Syntax Literate: Jurnal Ilmiah Indonesia p–ISSN: 2541-0849
e-ISSN: 2548-1398
Vol. 10, No. 1, Januari 2025

ASTRONOMICAL IMAGE DENOISING USING ATTENTIONGAN

Faishal Zaka Naufal[1], Muhammad Febrian Rachmadi[2], Adila Alfa Krisnadhi[3]
University Indonesia, Indonesia[1,2,3]
Email: faishal.zaka@ui.ac.id[1], febrian.rachmadi@cs.ui.ac.id[2], adila@cs.ui.ac.id[3]

Abstract

Denoising astronomical images is a significant challenge in the field of astronomical data processing. Image data acquired from astronomical sources typically contains noise from various sources. The study aims to investigate the denoising of astronomical images using an image-to-image translation approach with AttentionGAN method. This method combines attention-guided techniques with a Generative Adversarial Network (GAN) model to improve the quality of noisy astronomical images. Attention-guided technique allows the model to learn the most important features of the image and guide the image generation process. This approach has been tested on several images in different domains, each with varying levels of noise. The results shows that AttentionGAN method produces denoised images with better and sharper quality than several other denoising methods. Two databases, The Panoramic Survey Telescope and Rapid Response System (PAN-STARRS) and the Sloan Digital Sky Survey (SDSS), were used in this research. Images acquired from PAN-STARRS contain noise, while images acquired from SDSS are clean. Overall, this research contributes to improving the quality of astronomical images by demonstrating the effectiveness of the AttentionGAN method in denoising noisy astronomical images. We employed denoising techniques using CycleGAN and AttentionGAN and evaluated them using metrics such as PSNR, SSIM, and FID. The analysis showed that the AttentionGAN model outperformed CycleGAN. We also conducted ablation studies to further investigate the components of the AttentionGAN model. This study provides a foundation for future research in the field of astronomical data processing, which has the potential to enhance image quality.

Keywords: AttentionGAN, Denoising, Astronomical Image

Introduction

Astronomical images captured with telescopes play a crucial role in studying celestial objects. Astronomers use these images to analyze stars and other astronomical objects. Astronomical images are captured using high-resolution Charge-Coupled Device (CCD) sensors. CCD sensors work by converting light into electrical signals, resulting in digital images. However, the use of CCD sensors can introduce noise which originate from various sources including background sky or photon noise caused by low-energy sources (Flamary, 2017), atmospheric distortions when capturing images from the Earth's surface (Schawinski et al., 2017), and other environmental factors or interferences. The amount of astronomical data collected within a year can reach the scale of Terabytes (Hao-ran et al., 2017).

Denoising is crucial in distinguishing celestial objects from noise during analysis, making it an essential post-processing step in astronomy. One of the denoising methods is Bayesian least squares - Gaussian scale mixtures (BLS-GSM). However, this method is not suitable for astronomical images (Burger et al., 2011). Other conventional methods such as Gaussian filtering have also been used in denoising astronomical images, but they

often come with a trade-off of producing more blurred images (Laine et al., 2021). These methods also frequently fail to preserve edges, which are crucial components in the analysis process. When dealing with astronomical images, it is essential to retain edges while simultaneously removing noise and preserving the texture patterns of the images (Misra et al., 2018). In this regard, machine learning approaches have shown great potential in denoising by achieving high performance and avoiding blur through direct learning from the data itself (Laine et al., 2021).

Machine learning-based denoising approaches include the use of Convolutional Neural Networks (CNNs). CNNs can also be applied to tasks such as style transfer (Chen et al., 2017), super-resolution (Zhang et al., 2019), and segmentation (Carreira et al., 2012). For image-to-image mapping, networks can consist of convolutional (Flamary, 2017) and deconvolutional layers (Wang et al., 2023). Other methods that can be employed include Generative Adversarial Networks (GANs) (Lin et al., 2021) and U-net (Vojtekova et al., 2021).

Lin et al. (2020) stated that denoising research can be conducted using image-to-image mapping or image translation approaches. Denoising using image-to-image mapping, such as GAN, has proven to restore missing features in degraded images (Schawinski et al., 2017). Recent studies on denoising astronomical images have utilized paired input images, consisting of clean, noise-free images as ground truth and the same images with additional noise. Examples of such research include Schawinski et al. (2017) which added artificial degradation to the ground truth images, Vojtekova et al. (2020) which used images with shorter exposure times, and Flamary (2017) which introduced Gaussian noise. These studies employed GAN and U-Net methods.

Prior study also uses GAN to do denoising. Schawinski et al. (2017) employed GAN to restore features in astronomical images by denoising images with artificial noise. Fussell and Moews (2019) utilized chained GANs to generate high-resolution synthetic galaxy images, using a deep convolutional GAN and StackGAN architecture. Mustafa et al. (2019) applied CosmoGAN in the task of generating weak gravitational lensing convergence maps, which are a visual representation of the gravitational lensing effect on galaxy distribution. By altering observed galaxy density through gravitational lensing effect, scientists can measure the distribution of matter in the universe. The successful applications of GAN in astronomy research suggest a promising future in exploring its potential for other astronomical tasks.

In astronomy, obtaining paired images can be challenging as survey telescopes are often located in different positions on the Earth's surface, resulting in varying environmental and sky conditions. Therefore, research on image translation for unpaired astronomical images is crucial for efficient research. Denoising with unpaired images is relatively uncommon, but studies like GAN-based image translation (Lin et al., 2021) are emerging in astronomy. As a result, denoising methods for unpaired images are essential in processing astronomical data. One such method is the use of GANs with the cycle-consistency loss principle in image translation, as demonstrated by Lin et al. (2021). The study utilizes GANs to translate astronomical images across different domains by reconstructing noise in the target domain. In other domains outside of astronomy, image translation research has frequently employed the cycle-consistency loss principle. Tang et al. (2019) point out a limitation of Cycle-GAN, where certain feature elements receive too much attention compared to other focus and main features that cannot be captured quickly. Attention modules have been added to GAN architectures several times before. Jiang et al. (2022) used Cycle-GAN with the addition of a self-attention mechanism.

Compared to several standard Cycle-GAN methods, the proposed method has been shown to produce better results in terms of image quality, visual appeal, and noise reduction.

The use of CycleGAN has proven to be effective in denoising without the need for paired images. However, it tends to overemphasize certain features in the images, leading to distortions. This limitation affects the capture of essential features in a short period. To address this, attention modules have been added to the generator in CycleGAN, as demonstrated in the research by Tang et al. (2019). Tang et al. (2023) used Attention-Guided Generative Adversarial Networks (GANs) for unpaired image-to-image translation. It utilizes an attention-guided technique that allows the GAN model to learn more important features in the input and output images. An attention map is generated from the input image and used to guide the process of generating the output image. Compared to other GAN methods, Attention-GAN produces more realistic and higher quality images with better results. Evaluations of these attention-based methods have shown superior performance compared to traditional CycleGAN in various image translation tasks. However, AttentionGAN has not been applied to the domain of astronomical images and has primarily been used for denoising in general contexts. Inspired by the work of Tang et al. (2023), we aim to explore denoising through image translation using the AttentionGAN method in the domain of astronomical images.

Research Methods

CycleGAN

CycleGAN is a deep learning algorithm used for image domain transfer without requiring paired training data (Zhu et al., 2017). It uses two Generative and two Discriminative models to convert between two different image domains. One of CycleGAN's main features is its ability to convert between vastly different domains, such as from horse to zebra images or from summer to winter images. The algorithm achieves this by using Unsupervised learning, allowing the model to learn feature representations from both domains without the need for paired data. CycleGAN learns the mapping function between two domains, X and Y.

Compared to standard GAN, Cycle-GAN has two main improvements (Jiang et al., 2022). The first improvement is on the input received by the Generator. In standard GAN, the input is random noise, which means that only random images can be generated and the quality of the generated image cannot be controlled. The second improvement is on the required training data. GAN requires paired training data, which is difficult to obtain in some cases (Jiang et al., 2022). Unlike GAN, Cycle-GAN does not require paired images in the training process. The principle of Cycle-GAN is based on the cycle-consistency loss using the loss function of GAN. The adversarial loss controls the generated images to approach the target images, and the cycle-consistency loss is used to maintain the content structure of the input image and the feature of the target image. When generating an image, Cycle-GAN is trained to find potential connections between multiple feature domains and thus transform it to the relevant domain based on the input image (Zhu et al., 2020).

CycleGAN consists of two generators, namely G and F, and two discriminators, namely D_X and D_Y. The generator G_X learns the mapping from domain X to domain Y, while the generator G_Y learns the mapping from domain Y to domain X. The discriminator D_Y is used to determine whether the input image is generated by G_X, i.e., $G_X(x)$, and provides feedback to the generator G_X. Similarly, the discriminator D_X is used to

determine whether the input image is generated by G_Y, i.e., $G_Y(y)$, and provides feedback to the generator G_Y. The architecture of CycleGAN can be seen in Figure 1.

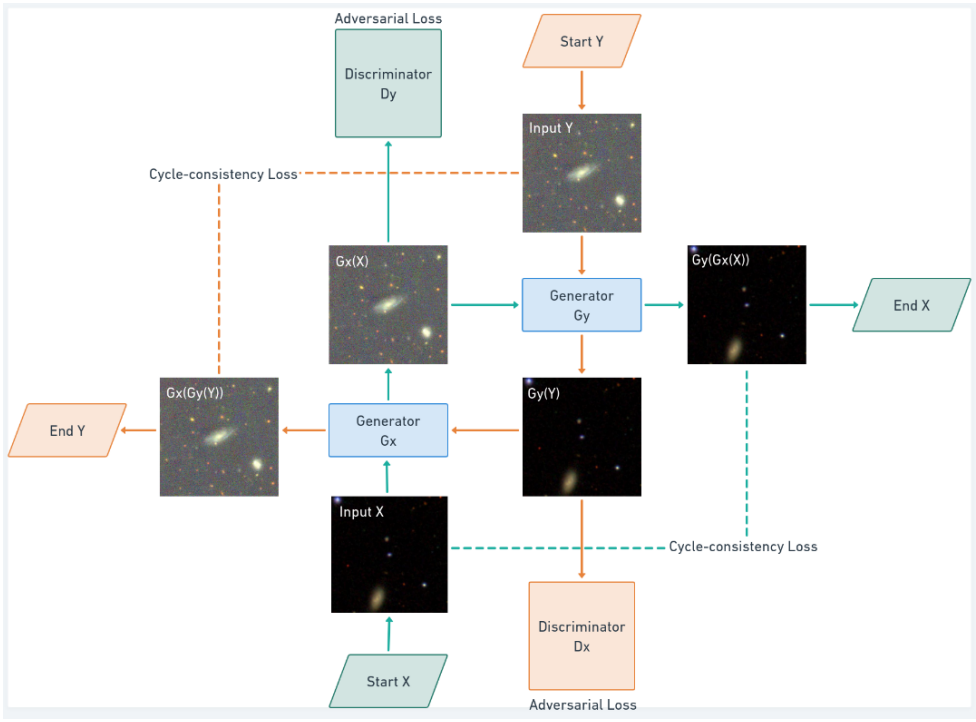


Figure 1. The diagram illustrates the structure of the CycleGAN network

The objective function of CycleGAN is divided into two parts: the first part is the adversarial loss from the GAN with two mapping functions, and the second part is the cycle-consistency loss. The adversarial loss of Generator G_X and discriminator D_Y is formulated by equation 1.

$$L_{GAN}(G_X, D_Y, X, Y) = E_{y \sim P_{data}(y)}[\log D_Y(y)] + E_{x \sim Pdata(x)}\left[\log\left(1 - D_Y(G_X(x))\right)\right] \quad (1)$$

In equation 1, G_X is the generator, D_Y is the discriminator, X and Y are the domains corresponding to $x \in X$ and $y \in Y$. $p_{data}(x)$ represents the distribution of the dataset X, and $E_{\_x \sim Pdata(x)}$ represents the mean value in the case of $p_{data}(x)$ on x. $p_{data}(y)$ represents the distribution of the dataset Y, and $E_{\_y \sim Pdata(y)}$ represents the mean value in the case of $p_{\_data}(y)$ on y. $G_X(x)$ is the generator G_X that generates an image from domain X to domain Y. The adversarial loss of Generator G_Y and discriminator D_X is formulated by equation (2).

$$L_{GAN}(G_Y, D_X, X, Y) = E_{x \sim P_{data}(x)}[\log D_X(x)] + E_{y \sim Pdata(y)}\left[\log\left(1 - D_X(G_Y(y))\right)\right] \quad (2)$$

In equation (2), G_Y is the generator, D_X is the discriminator, X and Y are the two bidirectional domains with $x \in X$ and $y \in Y$. $p_{data}(x)$ represents the distribution of the dataset X, and $E_{x \sim Pdata(x)}$ represents the mean value in the case of $p_{data}(x)$ at x. $p_{data}(y)$ represents the distribution of the dataset Y, and $E_{y \sim Pdata(y)}$ represents the mean value in the case of $p_{data}(y)$ at y. $G_Y(x)$ is the generator G that generates an image from domain X to domain Y.

In the training process, CycleGAN uses a loss function to measure the error between the generated images by the generator and the original images. This algorithm employs a cycle consistency loss technique, which allows the model to produce images that can be translated back to their original domain. By using these techniques, CycleGAN can generate images with relatively high quality in the cross-domain conversion process and has been applied in various applications, such as satellite image-to-map image conversion and artistic style transfer between different domains.

In CycleGAN, to address the inconsistency between the output distribution and the target distribution, a loss function called cycle-consistency loss is proposed to enforce the relationship between the generator input and output. The formulation of the cycle-consistency loss in CycleGAN is given by equation (3).

$$L_{cyc}(G_X, G_Y) = E_{x\sim P_{data}(x)}\left[\left|\left|G_Y(G_X(x)) - x\right|\right|_1\right] + E_{y\sim Pdata(y)} \times \left[\left|\left|G_X(G_Y(y)) - y\right|\right|_1\right] \quad (3)$$

In equation (3), G_X and G_Y are the generators, X and Y are the image domains with $x \in X$ and $y \in Y$. $p_{data}(x)$ represents the distribution of the dataset X, and $E_{x\sim Pdata(x)}$ represents the mean value in the case of $p_{data}(x)$ when x is sampled. $p_{data}(y)$ represents the distribution of the dataset Y, and $E_{y\sim Pdata(y)}$ represents the mean value in the case of $p_{data}(y)$ when y is sampled. $G_X(G_Y(y))$ is the generator G_X that takes an image from domain Y generated by the generator $G_Y(y)$ and produces an image in domain X. The result of this generator is an image in domain X. $G_Y(G_X(x))$ is the generator G_Y that takes an image from domain X generated by the generator $G_X(x)$ and produces an image in domain Y. The result of this generator is an image in domain Y. The final loss function in CycleGAN is formulated in equation (4)

$$L_{(G_X,G_Y,D_X,D_Y)} = L_{GAN}(G_X, G_Y, D_X, D_Y) + L_{cyc}(G_X, G_Y) \quad (4)$$

AttentionGAN

AttentionGAN (Tang et al., 2023) follows the principle of cycle-consistency loss with two generators, G and F, and two discriminators as shown in Figure 2. It consists of two sub-networks each, responsible for generating attention masks and content masks. G consists of a parameter-sharing encoder G_E, an attention mask generator G_A, and a content mask generator G_C. G_E aims to extract low-level and high-level feature representations. G_C aims to generate multiple intermediate content masks. G_A attempts to generate multiple attention masks. Both generators have their own network parameters and do not interfere with each other. The attention mask generator G_A aims to generate n - 1 foreground attention masks $\mathrm{A}_{\mathrm{y}}^{\mathrm{f}}$ and one background attention mask $\mathrm{A}_{\mathrm{y}}^{\mathrm{b}}$. By doing so, the network can simultaneously learn to generate new focused objects while maintaining the background in the input image.

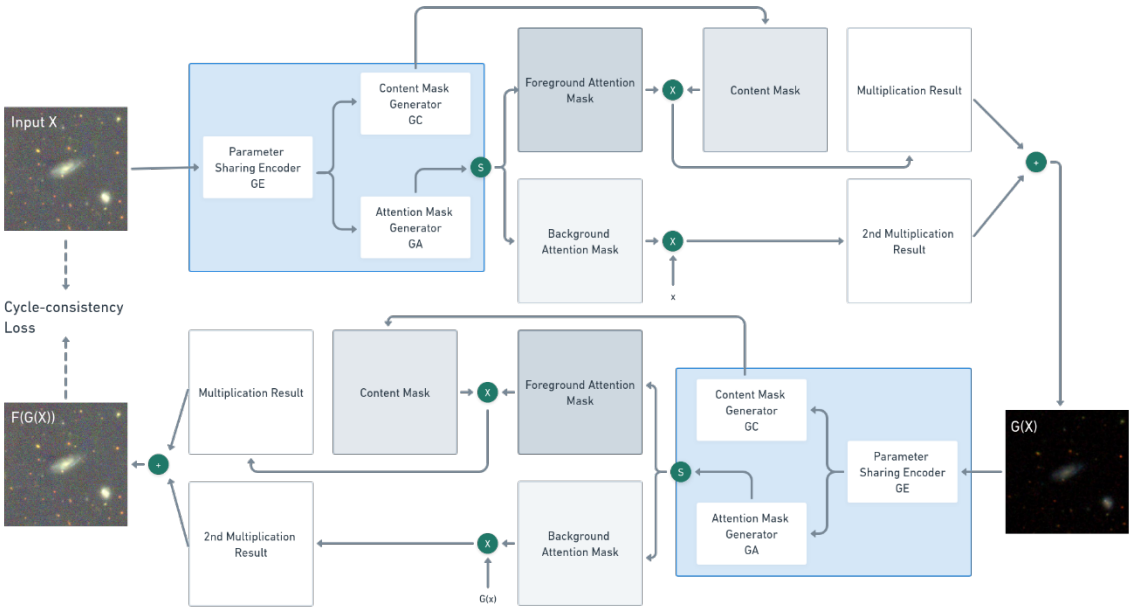


Figure 2. The diagram illustrates the structure of the AttentionGAN generator network

Feature map *m* extracted by G_E is consumed by the generator G_C to generate *n* - 1 content masks C_y^f , followed by the activation function *Tanh(.)*. This process is formulated by equation (5).

$$C_y^f = Tanh\left(mW_C^f + b_C^f\right) \qquad (5)$$

where $f = 1, \ldots, n - 1$, and the convolution operation is performed with convolution filters $W_C^f, b_{C\,f=1}^{n-1}$ t. Thus, the *n*-1 content masks and the input image *x* can be considered as parts of the candidate output image.

On the other hand, the feature map *m* is consumed by a group of filters $\mathrm{W_A^f, b_{Af=1}^{n-1}}$ to generate the corresponding *n* attention masks. This process is formulated by equation (6).

$$A_y = \mathrm{Softmax}\left(mW_A^f + b_A^f\right) \qquad (6)$$

where $f = 1, \ldots, n$, and *Softmax(.)* is the channel-wise softmax function for normalization. The attention masks $\mathrm{A_y}$ are then divided into *n* - 1 foreground attention masks $\mathrm{A_y^f}$ and one background attention mask A_y^b along the channel dimension. It should be noted that the background attention mask and the *n* - 1 generated foreground attention masks are complementary to each other, but the *n* - 1 foreground attention masks generated do not complement each other.

All attention masks will be multiplied by the content masks to obtain the final target output. This process is formulated by equation (7).

$$G(x) = \sum_{f=1}^{n-1}\left(C_y^f * A_y^f\right) + x * A_y^b \qquad (7)$$

The formulation of the generator *F* and the input image *y* can be expressed by equation (8).

$$F(y) = \sum_{f=1}^{n-1}\left(C_x^f * A_x^f\right) + y * A_x^b \qquad (8)$$

where a total of *n* attention masks are also generated by the channel-wise softmax activation function for normalization.

After *G(x)* generated by the generator *G*, it needs to be mapped back to the original domain to reduce the possible mappings that the model can perform. AttentionGAN has another generator, which is *F*. It has a similar structure to the generator *G* and also consists of three sub-networks: a parameter sharing encoder F_E, attention mask generator F_A, and content mask generator F_C. F_C will attempt to generate *n* - 1 content masks, and F_A will attempt to generate *n* background attention mask and foreground attention masks. Then, these masks are combined so that the image *G(x)* can reconstruct the original input image *x*. This process is formulated by equation (9).

$$F\big(G(x)\big) = \sum_{f=1}^{n-1}\big(C_x^f * A_x^f\big) + G(x) * A_x^b \qquad (9)$$

where the final image *F(G(x))* will be similar to the original image *x*. Image *y* is formulated by equation (10).

$$G\big(F(y)\big) = \sum_{f=1}^{n-1}\big(C_y^f * A_y^f\big) + F(y) * A_y^b \qquad (10)$$

where the final image *G(F(y))* will be similar to the original image *y*. Optimization function of AttentionGAN formulized by equation (11).

$$\mathcal{L} = \mathcal{L}_{GAN} + \lambda_{cycle} * \mathcal{L}_{cycle} + \lambda_{id} * \mathcal{L}_{id} \qquad (11)$$

where $\mathcal{L}_{GAN}$, $\mathcal{L}_{cycle}$ and $\mathcal{L}_{id}$ is GAN, cycle-consistency and identity preserving loss. λ_{cycle} and λ_{id} are parameter controlling relative relationship between every term in the model.

Evaluation Metrics

To assess the performance of the researched model, a qualitative evaluation is conducted using the Peak Signal to Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Fréchet Inception Distance (FID). PSNR is calculated by equation (12)

$$PSNR = 10\log_{10}\left(\frac{L_{I_{denoised}}^2}{MSE}\right) \qquad (12)$$

While MSE calculated by equation (13)

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I_{denoised}(i,j) - I_{truth}(i,j)]^2 \qquad (13)$$

where *m* and *n* represent the dimensions of the images $I_{denoised}$ and I_{truth}, respectively. $I_{denoised}$ is the denoised image, I_{truth} is the ground truth image, and $L_{denoised}$ is the maximum gray level value of the denoised image $I_{denoised}$. The Structural Similarity Index (SSIM) is calculated using the equation (14)

$$SSIM(i,j) = \big[l(i,j)^{\alpha} \cdot c(i,j)^{\beta} \cdot s(i,j)^{\gamma}\big] \qquad (14)$$

where *s* represents the structure, *c* represents the contrast, and *l* represents the luminance. Structure, contrast, and luminance.

Evaluation is also performed using FID (Fréchet Inception Distance). FID is an evaluation metric used to assess the quality of images generated by generative models. FID measures the distance between the distributions of real and generated images in the feature space, which is calculated using a pre-trained Inception-V3 classification network on the ImageNet dataset. In the field of astronomy, ImageNet weights are commonly used as pre-trained models for classifying celestial objects. Studies such as Martinazzo et al. (2020) and Farrens et al. (2022) used ImageNet weights for star/galaxy object classification using CNN architectures and compared models with fine-tuning and models without fine-tuning. From these studies, it is known that models with ImageNet weights can still extract features from astronomical images. FID is calculated using the equation (15).

$$FID = |\mu - \mu_w|^2 + tr\left(\Sigma + \Sigma_w - 2(\Sigma\Sigma_w)^{\frac{1}{2}}\right) \quad (15)$$

where $N(\mu, \Sigma)$ is the multivariate normal distribution estimated using Inception V3 (Szegedy et al., 2016) calculated on real images, and $N(\mu_w, \Sigma_w)$ is the multivariate normal distribution calculated using Inception V3 on generated images. Qualitative evaluation is also conducted by observing and comparing the denoised images obtained using CycleGAN and AttentionGAN with the ground truth images. Qualitative analysis is also performed to check for missing objects.

Results and Discussiom

Dataset

The data acquisition process begins with obtaining 700 celestial object coordinates. Out of these 700 coordinates, they are then divided into two sets, each containing 350 sky coordinates. These two sets are utilized to acquire data from two data sources, namely the Sloan Digital Sky Survey (SDSS) and the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS). The SDSS images represent clean image data, while the Pan-STARRS images contains noises. This is done to ensure that the acquired images are unpaired. These two sets of images will then serve as the training data and input into the CycleGAN and AttentionGAN models.

As part of the evaluation process, a total of 150 distinct sky coordinate data points from the training data are used. These data points are utilized to obtain noised images sourced from Pan-STARRS, as well as ground truth images sourced from SDSS. Therefore, during the evaluation process, the generated images from Pan-STARRS will be compared to the corresponding images from SDSS with the same sky coordinates, ensuring that the displayed objects are also identical.

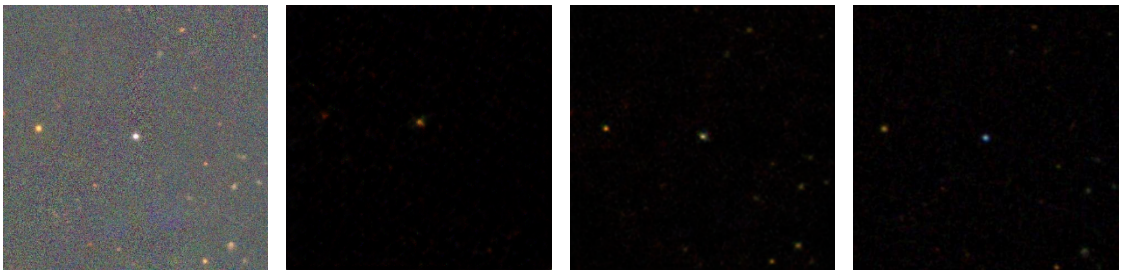
Evaluation using PSNR, SSIM and FID

A total of 700 data points are used for training, consisting of 350 SDSS images and 350 PANSTARRS images. Subsequently, the testing data is obtained by selecting 150 images from PANSTARRS as the noisy images and corresponding images from SDSS as the ground truth. The experiments for each method are repeated five times, and the average of each evaluation metric is calculated.

Table 1. Quantitative evaluation of the denoised image results using CycleGAN and AttentionGAN

Images	PSNR	SSIM	FID
SDSS (clean)	27.696	0.037	163.166
PanSTARRS (noised)	100	1.0	12.924
CycleGAN result	34.554	0.616	133.672
AttentionGAN result	**34.725**	**0.661**	**49.697**

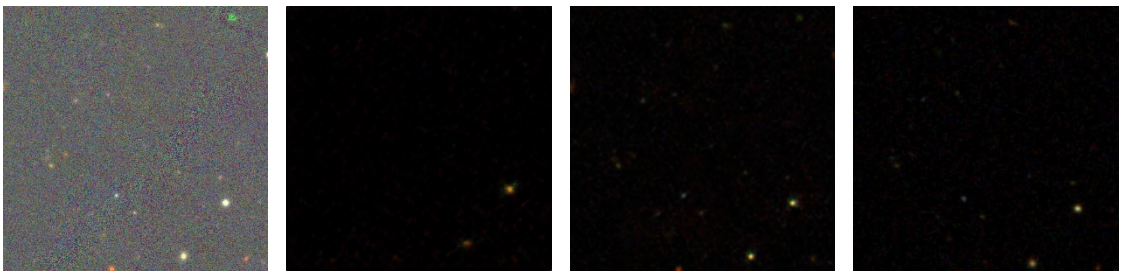
The PSNR evaluation metric is calculated by comparing the denoised images produced by both methods (CycleGAN and AttentionGAN) and comparing them to the clean images. The PSNR values of the noisy and clean images are also computed to measure the level of difference. Compared to CycleGAN, AttentionGAN has a higher average PSNR value of **34.725**, while CycleGAN has an average PSNR value of **34.554**. When considering individual images, AttentionGAN achieves the highest PSNR value of **36.821**, as shown in Figure 3(c). For the same source image, CycleGAN yields a PSNR value of 36.184, as indicated in Figure 3(b).



(a) Noised image **(b) CycleGAN result** **(c) AttentionGAN result** **(d) Ground truth**

Figure 3. Image that has highest PSNR value using AttentionGAN, compared with original noised image, CycleGAN result and ground truth image

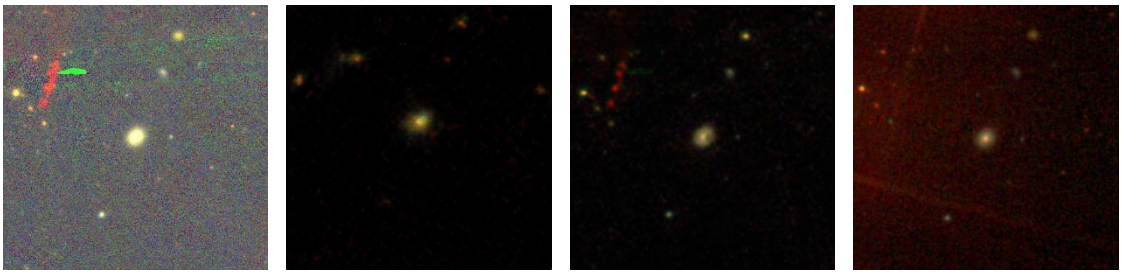
Using qualitative analysis, the images generated by the AttentionGAN method exhibit better quality. In the denoised image produced by AttentionGAN, five objects are detected, with two bright objects in the middle and left, and three dimmer objects in the bottom-right. When compared to the ground truth image, the number of detected objects appears to be the same. In contrast, when comparing it to the denoised image using CycleGAN, only two objects are detected in the middle and left. This suggests that there are celestial objects missing in the denoised image produced by CycleGAN.



(a) Noised image **(b) CycleGAN result** **(c) AttentionGAN result** **(d) Ground truth**

Figure 4. Image that has highest PSNR value using CycleGAN, compared with original noised image, AttentionGAN result and ground truth image

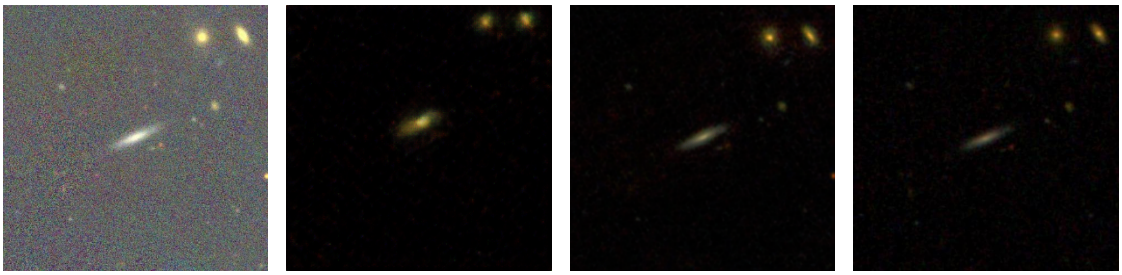
The highest PSNR value obtained with the CycleGAN method is **36.558**, as shown in Figure 4(b). Meanwhile, with the AttentionGAN method, a slightly lower PSNR value of 36.171 is achieved, as indicated in Figure 4(c). From a qualitative standpoint, two celestial objects are detected in the bottom-right region when compared to the ground truth image. Both the AttentionGAN and CycleGAN denoised images successfully remove noise without eliminating the visible celestial objects. However, there are positional changes between the images produced by the two methods. The largest positional differences are observed in the CycleGAN denoised image, while the positional changes in the AttentionGAN result are less pronounced.



(a) Noised image **(b) CycleGAN result** **(c) AttentionGAN result** **(d) Ground truth**

Figure 5. Image that has lowest PSNR value using both CycleGAN and AttentionGAN, compared with original noised image and ground truth image

The lowest PSNR values for CycleGAN and AttentionGAN in denoising are **29.888** and **30.641**, respectively, for the same source image as shown in Figure 5(b) and Figure 5(c). In the ground truth image, it can be observed that there are five visible celestial objects. Four objects form a straight line from the top-right to the bottom, while one object is located on the far left. In the AttentionGAN result, the visible celestial objects are preserved, but this method introduces new reddish artifacts on the left side. With the CycleGAN method, no new artifacts are apparent, but it only retains three visible celestial objects.



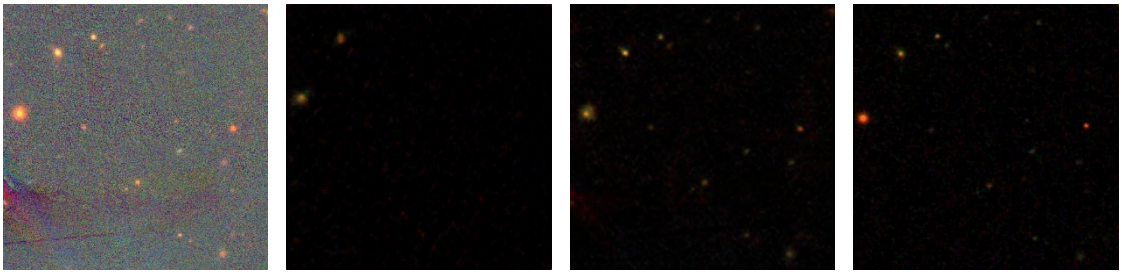
(a) Noised image **(b) CycleGAN result** **(c) AttentionGAN result** **(d) Ground truth**

Figure 6. Image that has highest SSIM value using AttentionGAN, compared with original noised image, CycleGAN result and ground truth image

Using the SSIM evaluation metric, the average value obtained during the denoising process with the CycleGAN method is **0.616**. This value is smaller compared to the AttentionGAN method, which has an average SSIM value of **0.661**. Considering individually, the highest SSIM value in the AttentionGAN method is **0.768**, as shown in Figure 6(c). For the same noisy image, the CycleGAN method yields an SSIM value of

0.618, as indicated in Figure 6(b). The smallest SSIM value in the AttentionGAN result is **0.542**, depicted in Figure 5(c).

Looking at the SSIM values for each image, the highest SSIM value of **0.691** is obtained with the CycleGAN method, as shown in Figure 7(b). With the AttentionGAN method, a relatively higher SSIM value of **0.740** is achieved, as indicated in Figure 7(c). The smallest SSIM value for CycleGAN is **0.311**, depicted in Figure 5(b).



(a) Noised image **(b) CycleGAN result** **(c) AttentionGAN result** **(d) Ground truth**

Figure 7. Image that has highest SSIM value using CycleGAN, compared with original noised image, AttentionGAN result and ground truth image

Using the FID evaluation metric, the FID score for the CycleGAN generated images is **133.673**. The FID score for the AttentionGAN generated images is better, with a value of **49.697**. The FID scores for the noisy and clean images are also calculated to assess the level of difference. For the noisy image, the FID score is 163.167, while for the clean image, the FID score is 12.924.

A t-test was conducted to analyze the significance of differences in multiple evaluation metrics based on the five conducted experiments. The t-test was performed with a significance level parameter α set to 0.05. The t-test was calculated involving the results of both the PSNR and SSIM metrics for the denoised images generated by CycleGAN and AttentionGAN. The obtained *p*-value for the PSNR metric is **0.087**, while the *p*-value for the SSIM metric is **0.117**. Considering that both *p*-values are greater than α, it can be concluded that the denoising results using both methods do not exhibit statistically significant differences, even though the denoising results with AttentionGAN are superior. Although the t-test results do not indicate a significant difference, the denoising technique using AttentionGAN is more capable of preserving celestial objects compared to CycleGAN. This is crucial for astronomical scientific analysis.

Ablation Study

We conducted an ablation analysis to examine the effects of component changes in the AttentionGAN model. In this case, changes were made to the generator component in generating denoised images, including: (1) removing the background mask component in the design (9f), (2) using only one foreground mask and one background mask (1f1b), and (3) using only one background mask without any foreground mask (1b). These three ablation processes were then analyzed using the evaluation metrics PSNR, SSIM, and FID. Furthermore, an error analysis was performed on all evaluation metrics. The evaluation values for the three ablation models are presented in Table 2.

Table 2. Quantitative evaluation of the AttentionGAN-Full model compared to the ablated models 9f, 1f1b, and 1b.

Model	PSNR (p-value)	SSIM (p-value)	FID (p-value)
AttentionGAN-Full	34.72	0.66	49.70
AttentionGAN-9f	**34.76 (0.80)**	**0.70 (0.12)**	**32.95 (0.01)**
AttentionGAN-1f1b	33.49 (0.14)	0.64 (0.69)	42.47 (0.09)
AttentionGAN-1b	27.63 (1.7×10^{-7})	0.03 (2.2×10^{-6})	94.73 (1.9×10^{-5})

From the first ablation result using the combination of 9f, there was an improvement in the average values of PSNR, SSIM, and FID. The average values of PSNR, SSIM, and FID for the ablated model 9f were 34.76, 0.70, and 32.95, respectively. The resulting t-test *p*-values between AttentionGAN-Full and AttentionGAN-9f were 0.80 for PSNR, 0.12 for SSIM, and 0.01 for FID. In the second ablation with the combination of 1f1b, the average values of PSNR, SSIM, and FID relatively decreased, with average values of 33.49, 0.64, and 42.47, respectively. The t-test *p*-values were 0.14 for PSNR, 0.69 for SSIM, and 0.09 for FID. Both results indicate that the absence of a background mask produced by the generator improves the quality of denoised images, albeit not significantly based on the *p*-values. The same applies to the reduction in the number of foreground masks in the denoised images.

In the third ablation, there was a significant decrease in the values of PSNR, SSIM, and FID. The average values of PSNR, SSIM, and FID resulting from the third ablation were 27.63, 0.03, and 94.73, respectively. These results indicate that the absence of a foreground mask has a significant impact on the denoising technique using the AttentionGAN method. Among the three ablated models, the AttentionGAN-9f model has better evaluation scores compared to the other ablated models. For each denoised image, the highest PSNR and SSIM values are 37.21 and 0.79, respectively. On the other hand, the lowest PSNR and SSIM values are 30.44 and 0.48, respectively.

Next, to examine the impact of different numbers of foreground masks used, an additional ablation study was conducted with the combinations of 4 foreground masks and 1 background mask (4f1b), as well as 7 foreground masks and 1 background mask (7f1b). The obtained results can be seen in Table 3.

Table 3. Quantitative evaluation of the AttentionGAN-Full model compared to the ablated models 1f1b, 4f1b, 7f1b.

Model	PSNR (p-value)	SSIM (p-value)	FID (p-value)
AttentionGAN-Full	34.72	0.66	49.70
AttentionGAN-1f1b	33.49 (0.14)	0.64 (0.69)	42.47 (0.09)
AttentionGAN-4f1b	**34.48 (0.36)**	**0.69 (0.41)**	**42.17 (0.11)**
AttentionGAN-7f1b	33.44 (0.08)	0.63 (0.45)	50.29 (0.91)

In Table 3, the best PSNR and SSIM values were obtained for the ablation model with a combination of 4 foreground masks and 1 background mask. These results were obtained after 5 experimental trials. However, these values are still lower compared to the ablation model with the combination \$9f\$ and the AttentionGAN-Full model. Based on the p-value for PSNR, there is no significant difference between the three ablation models.

Conclusion

In terms of the evaluation metrics PSNR, SSIM, and FID, the denoised astronomical images using AttentionGAN outperform the previously used model, CycleGAN. Additionally, after conducting t-test error analysis and ablation studies, it could be summarized that the differences in the average values of PSNR, SSIM, and FID between AttentionGAN and CycleGAN are considered statistically insignificant based on the p-value obtained from the t-test error analysis. Furthermore, the absence of background or foreground masks in the AttentionGAN generative model significantly affects the production of denoised images. Through the ablation study, it was found that the best average values of PSNR, SSIM, and FID are achieved when the generative model uses nine foreground masks. In terms of qualitative analysis, the AttentionGAN method performs better in preserving the position and shape of celestial objects. On the other hand, CycleGAN is more effective in reducing noise, albeit with the trade-off of changes in the position of celestial objects and the loss of some celestial objects.

BIBLIOGRAPHY

Burger, H. C., Schölkopf, B., & Harmeling, S. (2011). Removing noise from astronomical images using a pixel-specific noise model. *2011 IEEE International Conference on Computational Photography, ICCP 2011*. https://doi.org/10.1109/ICCPHOT.2011.5753128

Carreira, J., Caseiro, R., Batista, J., & Sminchisescu, C. (2012). Semantic segmentation with second-order pooling. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7578 LNCS*(PART 7). https://doi.org/10.1007/978-3-642-33786-4_32

Chen, Q., Xu, J., & Koltun, V. (2017). Fast Image Processing with Fully-Convolutional Networks. *Proceedings of the IEEE International Conference on Computer Vision*, *2017-October*. https://doi.org/10.1109/ICCV.2017.273

Farrens, S., Lacan, A., Guinot, A., & Vitorelli, A. Z. (2022). Deep transfer learning for blended source identification in galaxy survey data. *Astronomy and Astrophysics*, *657*. https://doi.org/10.1051/0004-6361/202141166

Flamary, R. (2017). Astronomical image reconstruction with convolutional neural networks. *25th European Signal Processing Conference, EUSIPCO 2017*, *2017-January*. https://doi.org/10.23919/EUSIPCO.2017.8081654

Fussell, L., & Moews, B. (2019). Forging new worlds: High-resolution synthetic galaxies with chained generative adversarial networks. *Monthly Notices of the Royal Astronomical Society*, *485*(3). https://doi.org/10.1093/mnras/stz602

Hao-ran, Q., Ji-ming, L., & Jun-yi, W. (2017). Stacked Denoising Autoencoders Applied to Star/Galaxy Classification. *Chinese Astronomy and Astrophysics*, *41*(2). https://doi.org/10.1016/j.chinastron.2017.04.009

Jiang, Z., Zhang, R., Guo, Y., Hu, M., He, L., Li, F., & Zhu, Z. (2022). Noise Interference Reduction in Vision Module of Intelligent Plant Cultivation Robot Using Better Cycle GAN. *IEEE Sensors Journal*, *22*(11). https://doi.org/10.1109/JSEN.2022.3164915

Laine, R. F., Jacquemet, G., & Krull, A. (2021). Imaging in focus: An introduction to denoising bioimages in the era of deep learning. In *International Journal of Biochemistry and Cell Biology* (Vol. 140). https://doi.org/10.1016/j.biocel.2021.106077

Lin, Q., Fouchez, D., & Pasquet, J. (2020). Galaxy image translation with semi-supervised noise-reconstructed generative adversarial networks. *Proceedings - International*

Conference on Pattern Recognition. https://doi.org/10.1109/ICPR48806.2021.9412143

Martinazzo, A., Espadoto, M., & Hirata, N. S. T. (2020). Self-supervised learning for astronomical image classification. *Proceedings - International Conference on Pattern Recognition*. https://doi.org/10.1109/ICPR48806.2021.9412911

Misra, D., Mishra, S., & Appasani, B. (2018). Advanced image processing for astronomical images. *ArXiv Preprint ArXiv:1812.09702*.

Mustafa, M., Bard, D., Bhimji, W., Lukić, Z., Al-Rfou, R., & Kratochvil, J. M. (2019). CosmoGAN: creating high-fidelity weak lensing convergence maps using Generative Adversarial Networks. *Computational Astrophysics and Cosmology*, *6*(1). https://doi.org/10.1186/s40668-019-0029-9

Schawinski, K., Zhang, C., Zhang, H., Fowler, L., & Santhanam, G. K. (2017). Generative adversarial networks recover features in astrophysical images of galaxies beyond the deconvolution limit. *Monthly Notices of the Royal Astronomical Society: Letters*, *467*(1). https://doi.org/10.1093/mnrasl/slx008

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016-December*. https://doi.org/10.1109/CVPR.2016.308

Tang, H., Liu, H., Xu, D., Torr, P. H. S., & Sebe, N. (2023). AttentionGAN: Unpaired Image-to-Image Translation Using Attention-Guided Generative Adversarial Networks. *IEEE Transactions on Neural Networks and Learning Systems*, *34*(4). https://doi.org/10.1109/TNNLS.2021.3105725

Tang, H., Xu, D., Sebe, N., & Yan, Y. (2019). Attention-Guided Generative Adversarial Networks for Unsupervised Image-to-Image Translation. *Proceedings of the International Joint Conference on Neural Networks*, *2019-July*. https://doi.org/10.1109/IJCNN.2019.8851881

Vojtekova, A., Lieu, M., Valtchanov, I., Altieri, B., Old, L., Chen, Q., & Hroch, F. (2021). Learning to denoise astronomical images with U-nets. *Monthly Notices of the Royal Astronomical Society*, *503*(3). https://doi.org/10.1093/mnras/staa3567

Wang, H., Sreejith, S., Lin, Y., Ramachandra, N., Slosar, A., & Yoo, S. (2023). Neural Network Based Point Spread Function Deconvolution For Astronomical Applications. *Open Journal of Astrophysics*, *6*. https://doi.org/10.21105/astro.2210.01666

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. *Proceedings of the IEEE International Conference on Computer Vision*, *2017-October*. https://doi.org/10.1109/ICCV.2017.244