Draft version September 24, 2025
Typeset using LaTeX twocolumn style in AASTeX7.0.1

Generative AI for image reconstruction in Intensity Interferometry: a first attempt

Km Nitu Rai,[1,2] Yuri van der Burg,[3] Soumen Basak,[1] Prasenjit Saha,[3] and Subrata Sarangi[4,5]

[1] *School of Physics, Indian Institute of Science Education and Research Thiruvananthapuram, Maruthamala PO, Vithura, Thiruvananthapuram 695551, Kerala, India.*

[2] *Aryabhatta Research Institute of Observational Sciences, Manora Peak, Nainital 263129, India.*

[3] *Physik-Institut, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland.*

[4] *School of Applied Sciences, Centurion University of Technology and Management, Odisha-752050, India.*

[5] *Visiting Associate, Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, Pune 411 007, Maharashtra, India.*

ABSTRACT

In the last few years Intensity Interferometry (II) has made significant strides in achieving high-precision resolution of stellar objects at optical wavelengths. Despite these advancements, phase retrieval remains a major challenge due to the nature of photon correlation. This paper explores the application of a conditional Generative Adversarial Network (cGAN) to tackle the problem of image reconstruction in Intensity Interferometry. This approach successfully reconstructs the shape, size, and brightness distribution of a fast-rotating star from sparsely sampled, spatial power spectrum of the source, corresponding to II with four telescopes. Although this particular example could also be addressed using parameter fitting, the results suggest that with larger arrays much more complicated systems could be reconstructed by applying machine-learning techniques to II.

1. INTRODUCTION

Intensity Interferometry (II) was first reported by Hanbury Brown and Twiss (HBT) during the 1950s (R. Hanbury Brown et al. 1954; R. Hanbury Brown & R. Q. Twiss 1956) as a "new type of interferometry" to measure stellar parameters such as angular diameter, orbits, and limb darkening coefficients. Later, theoretical results reported by R. Hanbury Brown et al. (1957, 1958), along with those of R. J. Glauber (1963) and others, demonstrated the deeper physical properties of photon correlations that lie at the core of II and laid the foundation for Quantum Optics (for textbook treatments see M. Leonard et al. 1995; E. Hecht 2002).

By the 1970s, with stellar parameter measurements of 32 stars in single and multiple star systems conducted by Hanbury Brown and his collaborators (R. Hanbury Brown et al. 1974) at the historic Narrabri Stellar Intensity Interferometer (NSII) in Australia, II had emerged as an alternative to the already established technique of Michelson Interferometry for measuring stellar parameters. Despite these significant achievements, the method did not gain widespread adoption in the ensuing decades, primarily due to the unavailability of sensitive photon detectors and advanced data analysis equipment.

More recently, proposals to utilize Imaging Atmospheric Cherenkov Telescope (IACT) facilities for conducting II observations of stars have emerged(J. Le Bohec 2006; P. D. Nuñez et al. 2010, 2012a; D. Dravins et al. 2013). It has been demonstrated that such observations could be carried out during bright, moonlit nights when γ-ray observations based on upper atmospheric Cherenkov showers were not feasible. This approach has the potential to enhance the scientific output of existing IACT facilities, and especially of the upcoming Cherenkov Telescope Array Observatory (CTAO). SII observations at VERITAS, MAGIC, and HESS are now being reported (e.g., A. Acharyya et al. 2024; S. Abe et al. 2024; N. Vogel et al. 2025). Simulations (e.g., K. N. Rai et al. 2021, 2022) have argued that recent advancements in photon detectors could be effective in achieving high-precision measurements of parameters for stellar objects.

Beyond measuring stellar diameters and other parameters of star systems, a fundamental goal of optical astronomy is to image stellar systems at high angular resolution. In the context of II, this involves reconstructing the source's image from the intensity correlations recorded by pairs of telescopes (light buckets) on the ground. However, because the primary observable in II is the electromagnetic field intensity rather than the field amplitude, the phase of the interferometric signal is lost. Since a complete reconstruction of a source's brightness distribution requires phase information, the challenge is to recover the phase of the signal.

Several theoretical and computational approaches for phase reconstruction with II have been proposed. H. Gamo (1963) introduced the concept of triple-intensity correlation, which M. L. Goldberger et al. (1963) subsequently applied in an experiment to observe scattered particles in microscopic systems. Sato conducted experiments to measure the diameter and phase of asymmetrical objects, suggesting that triple correlation could extend II to image stellar bodies (T. Sato et al. 1978, 1979, 1981). However, achieving a satisfactory signal-to-noise ratio (SNR) remained a significant challenge for this approach.

R. W. Gerchberg (1972) suggested an iterative method to determine the phase from the image and diffraction plane pictures. This method relies on accurate initial estimates and is vulnerable to slow convergence otherwise. J. Fienup (1982) introduced a Hybrid Input-Output algorithm that incorporates feedback mechanisms to improve convergence rate and robustness, particularly in noisy environments.

Later, R. Holmes et al. (2010) proposed an alternative method that utilizes the Cauchy-Riemann relations to reconstruct 1-D images. They also extended the approach to 2-D images across a range of signal-to-noise (SNR) values. This algorithm was applied to simulated data of stellar objects using II, considering both existing and forthcoming Imaging Cherenkov Telescope Arrays (IACTs) with a large number of telescopes (P. D. Nuñez et al. 2010, 2012a,b). However, this method faces challenges related to computational complexity when attempting to generalize to higher dimensions.

X. Li et al. (2014) suggested a flexible iterative Regularization method that incorporates prior information (e.g., sparsity, smoothness, or non-negativity) to reduce the ill-posedness of the phase retrieval problem. This method is more robust against noise and stabilizes the solution against artefacts and spurious solutions. Nevertheless, it faces challenges regarding the choice of the regularization parameter, computational complexity, and sensitivity to the initial guess.

The Transport-of-Intensity Equation (TIE) method is a non-interferometric technique first proposed by M. R. Teague (1983) that relates the intensity variations along the optical axis to the phase of the optical fields. This method enables phase retrieval from intensity measurements taken at multiple planes. J. Zhang et al. (2020) proposed a method to obtain a "universal solution" to the TIE by employing a "maximum intensity assumption", thereby converting the TIE into a Poisson equation which is then solved iteratively. More recently, C. Kirisits et al. (2024) have explored hybrid methods that combine the TIE with other equations, such as the Transport of Phase Equation (TPE). These approaches leverage the strengths of both equations to improve phase retrieval accuracy. This method is universally applicable, as it works for arbitrarily shaped apertures, handles non-uniform illumination, and accommodates inhomogeneous boundary conditions. It guarantees convergence, although the speed of convergence depends on the quality of the initial guess, and the final results are influenced by the boundary conditions.

With non-linearity built into their architecture, artificial neural networks (ANNs) empowered by deep learning methods are promising for exploring the task of reconstructing images of stellar objects from ground-based observations. Convolutional Neural Networks (CNNs), with their specialized architecture for processing two-dimensional datasets, are a natural choice for image processing tasks. In astronomical image reconstruction projects, a common challenge is that the interferometric data are typically undersampled as well as noisy. Therefore, the CNN architectures and deep learning methods employed must be capable of reliably learning both the global context of the training dataset and the local features within it. Among the various CNN architectures, U-Net models (O. Ronneberger et al. 2015) have proven successful in such tasks.

Furthermore, given that achieving a high signal-to-noise ratio (SNR) is often challenging in astronomical datasets, it is immensely beneficial if additional data can be generated using the available information from the observed sky density distribution and ground-based observations (II data, in our case) of the sources under investigation. Generative Adversarial Networks (GANs), introduced by I. Goodfellow et al. (2014), have been successful in such data augmentation tasks. Conditional GAN (cGAN) architectures, proposed by M. Mirza et al. (2014) and applied to a wide variety of datasets by P. Isola et al. (2017), leverage additional information about the images in the training datasets and have demonstrated remarkable robustness in image recovery across diverse data types.

In the astrophysical context, K. Schawinski et al. (2017) employed a GAN model to recover features — such as spiral arms, central bulges, and disk structures of galaxies — from noise-affected images. M. Mustafa et al. (2019) developed and customized a Deep Convolutional GAN, dubbed "CosmoGAN", capable of generating high-fidelity weak-lensing convergence maps of dark matter distribution that statistically reproduce real weak lensing structures. D. Coccomini et al. (2021) have successfully generated credible images of planets, nebulae, and galaxies using "lightweight" and "physics-uninformed" GANs to produce synthetic images of celes-

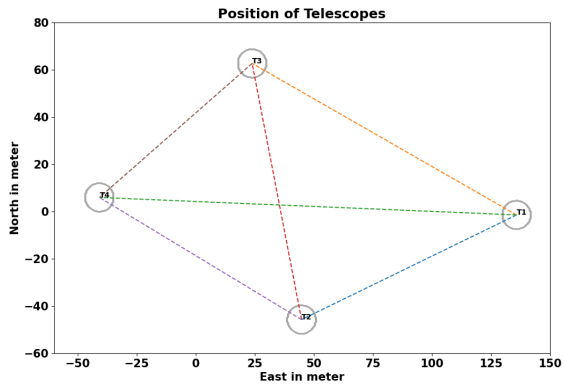

Figure 1. The telescope configuration with similar properties each used to simulate the signal for II observation.

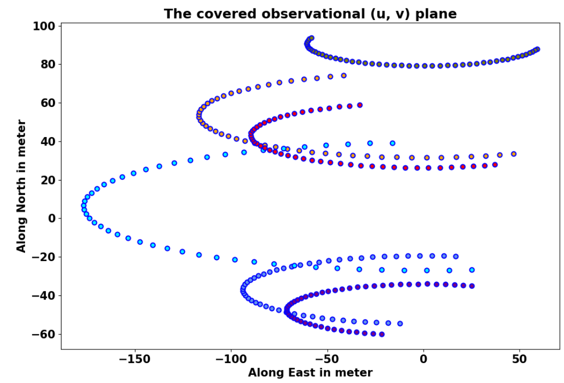

Figure 2. The tracks of the baselines provided by the four telescopes arranged in fig. 1 for one night of observation.

tial bodies. They also generated a “Hubble Deep Field-inspired” wide-view simulation of the universe.

In this paper, we propose a conditional Generative Adversarial Network (cGAN) model (following P. Isola et al. 2017) to reconstruct images of fast-rotating stars using their simulated Intensity Interferograms and simulated sky-intensity distributions as input data for training, testing, and validation. We consider four Imaging Cherenkov Telescope Arrays (IACTs) and simulate observation of a fast-rotating star. The image predicted by the trained GAN shows promising results in reconstructing the star’s shape and size. The reconstructed brightness distributions are then assessed using moments.

This paper is organized as follows. The next section discusses Intensity Interferometry, focusing on its signal and noise characteristics for fast-rotating stars along the Earth’s rotation. The following section introduces the GAN formulation and its structure. The fourth section details the parameter selection for training the GAN for image reconstruction. The fifth section presents the results of the trained GAN both visually and via image moments. Finally, the paper concludes with a discussion of the overall results.

2. INTENSITY INTERFEROMETRY (II) WITH IACT ARRAYS

This section presents a brief conceptual overview of how an array of telescopes is used to perform II observations, and explains the Signal-to-Noise Ratio (SNR) from these measurements.

2.1. *The signal for II*

As a simple example, let us consider a pair of IACTs pointed at a star. Suppose the two telescopes simultaneously measure the intensity of radiation $I_1(t)$ and $I_2(t)$, respectively. The signals from these detectors are cross-

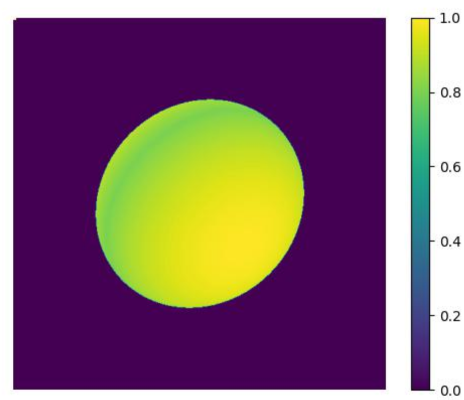

Figure 3. This figure shows the simulated fast rotating star. The brightness is highest at the poles and there is ~~gravitational~~ gravity darkening along the equator.

correlated and averaged over time, yielding the second order ($n = 2$) correlation of these intensities as (cf. V. A. Acciari et al. 2020; D. Dravins et al. 2013)

$$g^{(2)} = \frac{\langle I_1(t) \cdot I_2(t+\tau)\rangle}{\langle I_1(t)\rangle \cdot \langle I_2(t)\rangle} \tag{1}$$

where τ is the time delay between the telescopes. For spatially coherent and randomly polarized light, Eq. (1) reduces to the relation (sometimes called the Siegert relation, see e.g., V. A. Acciari et al. 2020).

$$g^{(2)} = 1 + \frac{\Delta f}{\Delta \nu} |V_{12}|^2 \tag{2}$$

where Δf is the electronic bandwidth of the photon detectors which measure the intensities and $\Delta\nu$ is the frequency bandwidth of the filters employed in the telescopes to observe the star. Values of $\Delta\nu \sim 1\,\mathrm{THz}$ and

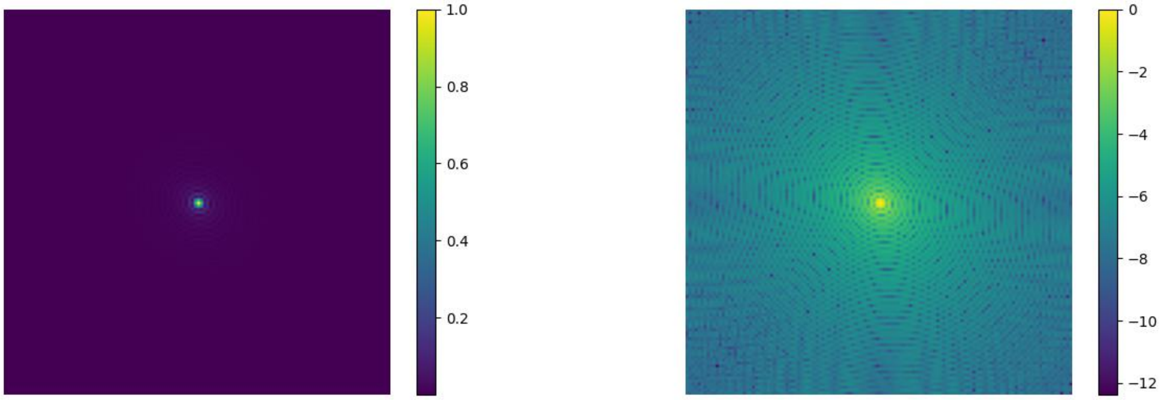

Figure 4. Absolute value of the two-dimensional Fast Fourier Transform of the source depicted in Fig. 3 in linear scale (left panel) and logarithmic scale (right panel). These figures represent the intensity interferometric (u, v) plane image of the source that would be obtained by an infinite number of baselines (or infinite number of telescopes observing the source). Both the linear and the logarithmic scales are normalized to the maximum intensity obtained at the centre of figures.

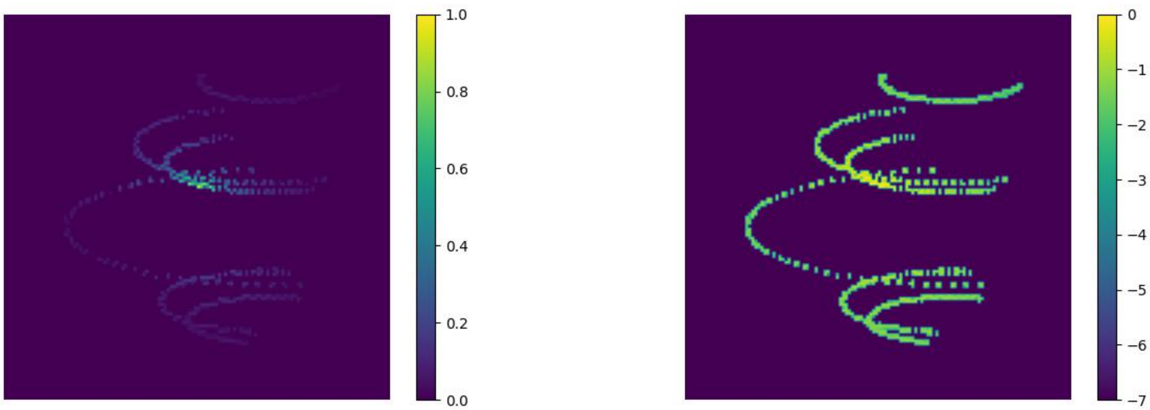

Figure 5. Absolute value of the two-dimensional Fast Fourier Transform of the source depicted in Fig. 3 and measured along the tracks shown in Fig. 2 covered by the baselines shown in Fig. 1. Both the left panel (in linear scale) and the right panel (in logarithmic scale) are normalized to the maximum pixel value in the respective figures. The panels of Fig. 5 reflect the sparse nature of the signal received by the realistic finite number of telescopes and baselines sampled from the full (u, v) plane signal space of Fig. 4.

$\Delta f \sim 1\,\mathrm{GHz}$ are typical of recent work. In Eq. (2), V_{12}, referred to as the complex visibility function, is the Fourier transform of the source brightness distribution. For a uniform disk source representing the star, it is given by

$$V_{12} = 2\frac{J_1(\pi\theta_D b)}{(\pi\theta_D b)} \tag{3}$$

where θ_D is the angular diameter of the star and b is the radial coordinate in the conventional interferometric $(u, v) = (x/\lambda, y/\lambda)$ plane, with λ representing the optical wavelength of the filter used for observation. It is evident from Eq. (3) that V_{12} contains information about the star's angular diameter. However, the phase information is lost since we measure only the absolute value $|V_{12}|^2$. In observational astronomy, the correlation is often expressed in terms of the normalized contrast, given by:

$$c = \frac{\langle (I_1(t) - \langle I_1 \rangle) \cdot (I_2(t+\tau) - \langle I_2 \rangle) \rangle}{\langle I_1(t) \rangle \cdot \langle I_2(t) \rangle} = g^{(2)} - 1 \tag{4}$$

where, $\langle I_1 \rangle$ and $\langle I_2 \rangle$ denote the mean intensities from the two telescopes. Therefore, the signal measured by the photon detectors in II, operating with an electronic bandwidth Δf within the optical bandwidth $\Delta\nu$ of the observational (filtered) radiation, is

$$c = g^{(2)} - 1 = \frac{\Delta f}{\Delta\nu}\left|V_{12}\right|^2 \tag{5}$$

with $|V_{12}|^2$ being a function of baseline $b = \sqrt{u^2+v^2}$ on the observational plane. This implies the strength of the signal would be enhanced if a larger number of baselines or pairs of telescopes are employed.

2.2. *The Signal-to-Noise Ratio for II*

The primary purpose of IACTs is to study high-energy gamma rays (with energy $E \geq 30$ GeV) arriving from cosmic sources, entering the Earth's atmosphere, and initiating Cherenkov showers in the upper atmosphere due to multiple scattering. These telescopes feature an array of mirrors that focus light onto a set of photomultiplier tubes (PMTs, see e.g., J. Aleksić et al. 2016). In the simulation model adopted here, we consider a set of four IACTs, each with similar properties. The positional configuration of these IACTs is shown in Fig. 1. The optical signal directed to a PMT is filtered using a spectral filter with a chosen mean observational wavelength λ and corresponding bandpass $\Delta\lambda$. The use of filters not only reduces background noise but also improves the signal quality and the efficiency of the PMTs. Filtering background skylight becomes even more significant in II observations, as, currently, these are carried out during full moon nights when the primary function of the IACTs (of observing Cherenkov Showers) is rendered infeasible. It is important to note that the light from the stellar source is focused on a PMT attached with each of the telescopes during II observations.

The significance of the signal can be expressed in terms of the signal-to-noise ratio (SNR), which depends on many factors. However, most importantly, it does not depend on the optical bandwidth $\Delta\nu$ of the radiation for a two-telescope correlation. The explanation for the independence of the SNR from $\Delta\nu$ is provided in several works (e.g., subsection 4.1 of K. N. Rai et al. 2021). The Signal-to-Noise is given by

$$SNR = A \cdot \alpha \cdot q \cdot n \cdot F^{-1} \cdot \sigma \cdot \sqrt{\frac{T\Delta f}{2}} \cdot \left|V_{12}\right|^2 \tag{6}$$

Here, A is the total mirror area, α is the quantum efficiency of the PMTs, q is the throughput of the remaining optics, and n is the differential photon flux from the source. The excess noise factor of the PMTs is represented by F, T denotes the observation time, and σ is the normalized spectral distribution of the light (including filters) (e.g., V. A. Acciari et al. 2020). The signal (S) and noise (N) can be inferred using eqns.5 and 6 as:

$$S = \frac{\Delta f}{\Delta\nu}\left|V_{12}\right|^2 \tag{7}$$

and

$$N = (A \cdot \alpha \cdot q \cdot n \cdot F \cdot \sigma \cdot \Delta\nu)^{-1}\sqrt{\frac{2\Delta f}{T}}. \tag{8}$$

While most of the parameters can be optimized with hardware, the only way to achieve a better SNR with fixed telescopes that is at the disposal of the astronomer is to increase the observation time T.

2.3. *Baseline considerations*

The measurement of the size of stellar objects via squared visibility depends on the distance between the telescopes, known as the baseline b.

$$\left|V_{12}(b)\right|^2 = \frac{c(b)}{c(0)} \tag{9}$$

For achieving a good SNR with a given telescope configuration, covering as much as possible of the interferometric plane is always desirable. If the source is at the zenith, the coordinates in the Fourier plane (u, v) are given by:

$$(u, v) = \frac{1}{\lambda}(b_E, b_N) \tag{10}$$

where b_E and b_N are the baselines expressed in east and north coordinates. However, of course sources can be anywhere on the sky, and the telescopes are stationary and may also have different relative altitudes b_A depending on the available terrain. Therefore, the Earth's rotation must be taken into account to cover the maximum observational plane using rotated baselines. For a given stellar source with declination δ and hour-angle h, as observed by telescopes at latitude l, equation (11) provides the rotated baselines for a given pair of telescopes (see e.g., eqs. 8–10 from S. Baumgartner et al. 2020).

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = R_x(\delta) \cdot R_y(h) \cdot R_x(-l) \begin{pmatrix} b_E \\ b_N \\ b_A \end{pmatrix} \tag{11}$$

Fig. 2 shows the track of six baselines generated from the telescopes (Fig. 1) due to the Earth's rotation. Since every pair of telescopes traces an ellipse in the Fourier plane, the total number of ellipses scales as

$$\mathcal{N} = \tfrac{1}{2} N_T \cdot (N_T - 1) \tag{12}$$

where N_T is the number of telescopes considered. As the number of baselines increases non-linearly, Intensity

Interferometry (II) benefits greatly from a large number of telescopes. The CTAO can offer many more baselines — D. Dravins et al. (2013) considered the telescope configurations then being planned and showed how it would provide a dense coverage of the interference plane.

2.4. *A Fictitious Fast Rotating Star: Our Test Case*

In our work presented here, we simulate a single fast-rotating star to test image reconstruction using a GAN. Fast rotation causes stars to adopt an oblate shape, flattening at the poles and bulging at the equator due to the stronger centrifugal force (e.g., H. Von Zeipel 1924; A. Maeder 1999). Fig. 3 shows an image qualitatively representing a fictitious fast-rotating star, with brightness distributed across its surface. The brightness is highest at the poles and lowest at the equator, a phenomenon known as gravity darkening (L. B. Lucy 1967). This effect was first observed through interferometric and spectroscopic data from the CHARA Array for the fast-rotating star Regulus H. A. McAlister et al. (2005). Fast-rotating stars are important test cases for understanding various astrophysical processes, including stellar evolution, internal structure, and dynamical behavior over time. It has not yet been observed using intensity interferometry, but oblateness has recently been measured (A. Archer et al. 2025) and gravity darkening is a natural next step.

Intensity Interferometry counts the photons arriving at the telescopes from the stellar object. The correlation of these photon arrivals at the telescopes yields the squared visibility Eq. (9), as explained in subsection 2.1. Fig. 4 shows the signal from the source shown in Fig. 3 using II, displayed in both linear and logarithmic scales. Point to note here is that this figure represents the signal from the source that would be recorded by an infinite number of baselines provided by an infinite number of telescopes on the interferometric plane. In practice, only a small part of this information is available (as seen Fig. 5), because one has a finite number of baselines corresponding to the finite number N_T of telescopes at our disposal and a limited observation schedule. We have simulated the II observation of the fictitious star by four telescopes (correlated with baselines as seen in Fig. 1) over one night. Using this modest amount of signal from one night's observation, we have trained a cGAN to construct the image of the source.

3. GENERATIVE ADVERSARIAL NETWORKS

Generative Adversarial Networks (GANs) were introduced by I. Goodfellow et al. (2014). The underlying concept is straightforward: it involves two competing networks. The first network, known as the Generator,

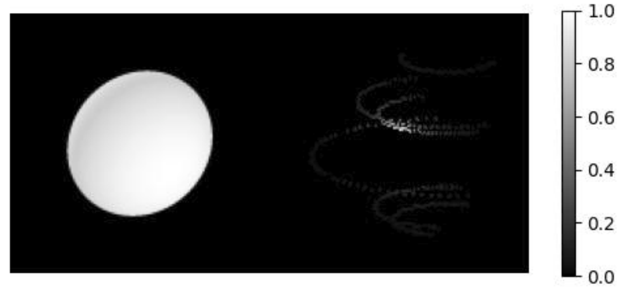

Figure 6. Merged image, which includes the original and the sparsely sampled Fourier plane. It is exactly what the GAN receives. The grey scale of the figure is normalized to the brightest pixel in the image.

produces new images based on an input image. These will be referred to as generated images. The second network, the Discriminator, attempts to distinguish between the generated image (predicted image) and the real image (ground truth).

Through the alternating training of these networks, the generated images gradually become indistinguishable from the real images. Essentially, this process constitutes a two-player min-max game — a classic problem in game theory. The original formulation of GANs is given by:

$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{\mathrm{data}}(x)}\left[\log D(x)\right] + \mathbb{E}_{z\sim p_z(z)}\left[\log\left(1 - D(G(z))\right)\right] \tag{13}$$

where $V(D,G)$ denotes the value function of the min-max game.

The objective is to learn the Generator's distribution, p_G, over the data x. We begin with input noise variables $p_z(z)$ and employ two perceptrons, $G(z;\theta_G)$ and $D(x;\theta_D)$, parameterized by θ_i with $i = G$ or D respectively. Here, $G(z)$ is a differentiable function that maps z to the data space x, while $D(x)$ represents the probability that x originates from real data. The problem can be reformulated as:

$$\max_D V(G,D) = \mathbb{E}_{x\sim p_{\mathrm{data}}}\left[\log D^*_G(x)\right] + \mathbb{E}_{x\sim p_G}\left[\log\left(1 - D^*_G(x)\right)\right] \tag{14}$$

where D^*_G denotes the optimum of the discriminator for a given fixed generator, as shown in equation (15). It can be demonstrated that the global optimum of equation (14) is achieved if and only if $p_G = p_{\mathrm{data}}$. Furthermore, if both G and D are allowed to reach their respective optima, then p_G converges to p_{data}. A more comprehensive discussion of the problem, including proofs, is provided in I. Goodfellow et al. (2014).

$$D^*_G(x) = \frac{p_{\mathrm{data}}(x)}{p_{\mathrm{data}}(x) + p_G(g)} \tag{15}$$

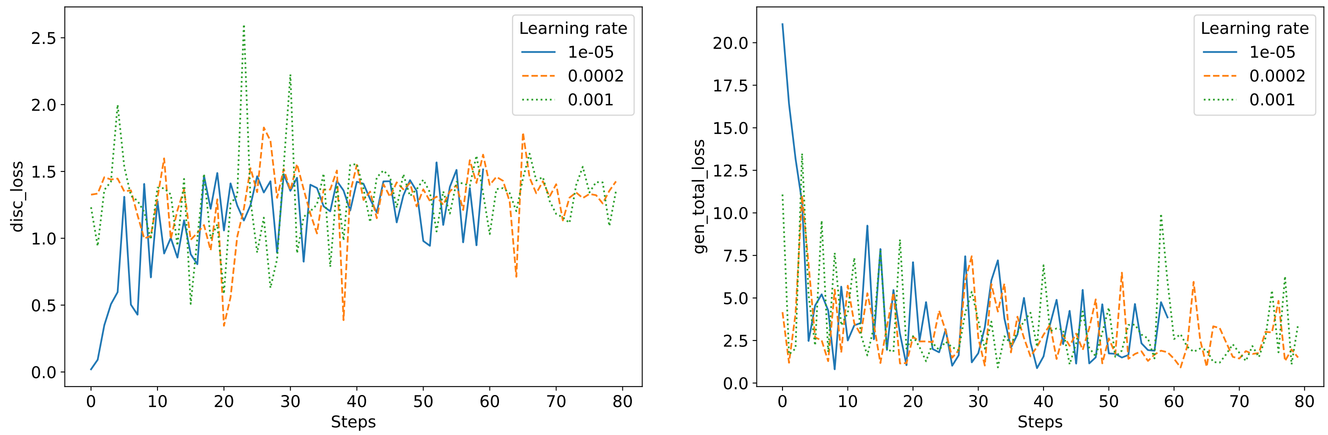

Figure 7. Discriminator and Generator losses for three different learning rates. The left panel and the right panel show the total discriminator loss and the total generator loss (Eq. 17) respectively. There is no significant difference, but these figures indicate that higher learning rates might render the training prone to outliers.

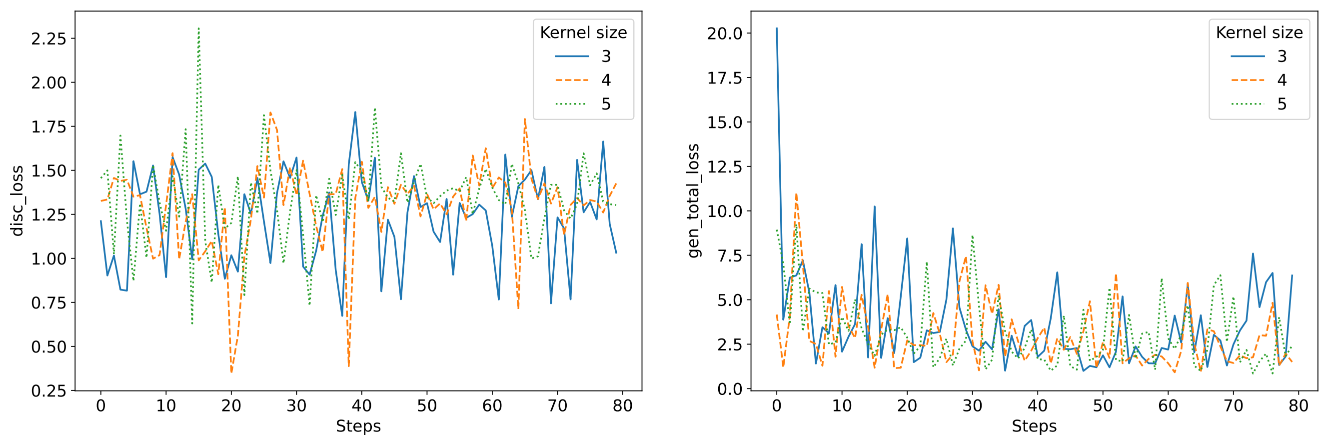

Figure 8. Discriminator and Generator losses (left and right panels respectively) for three different kernel sizes in the convolutional layers. Here, the smallest kernel size has many outliers, while the largest kernel size seems to be the most stable.

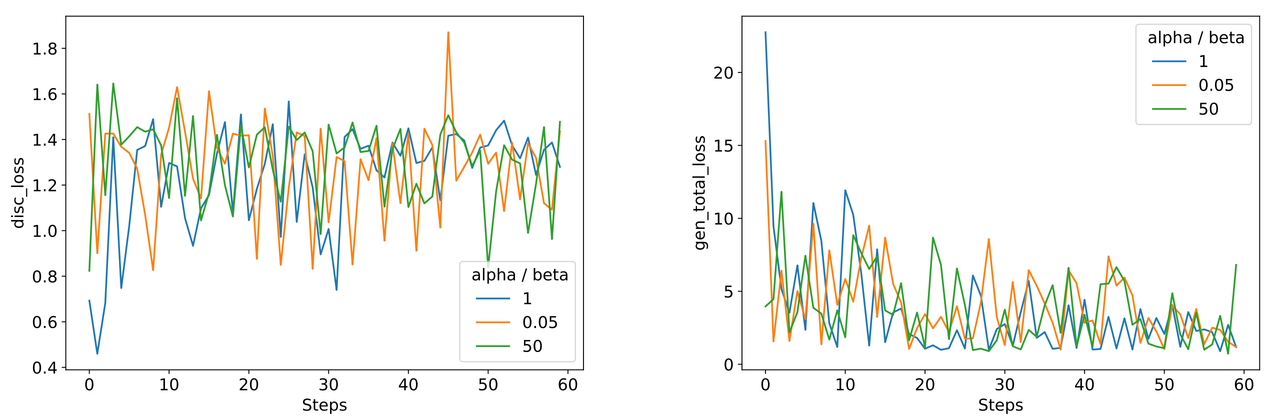

Figure 9. Effect of the Salt (alpha) and Pepper (beta) noise (explained in the text) introduced into the images. There is no significant effect of the alpha/beta ratio. These results are from training on 64×64-pixel images.

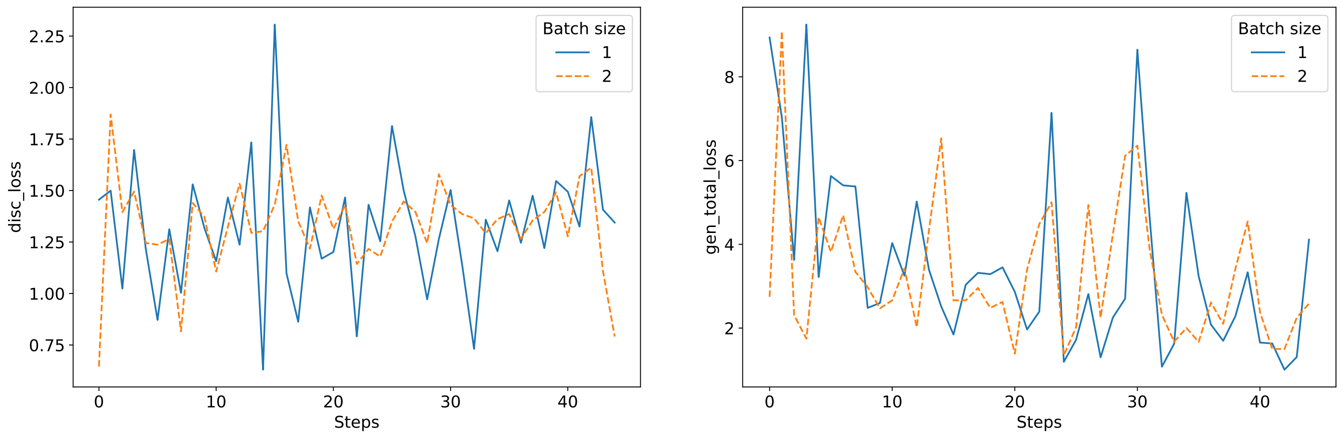

Figure 10. Loss functions for two different batch sizes. Large batch sizes seem to make the training more robust, but also increase training time significantly.

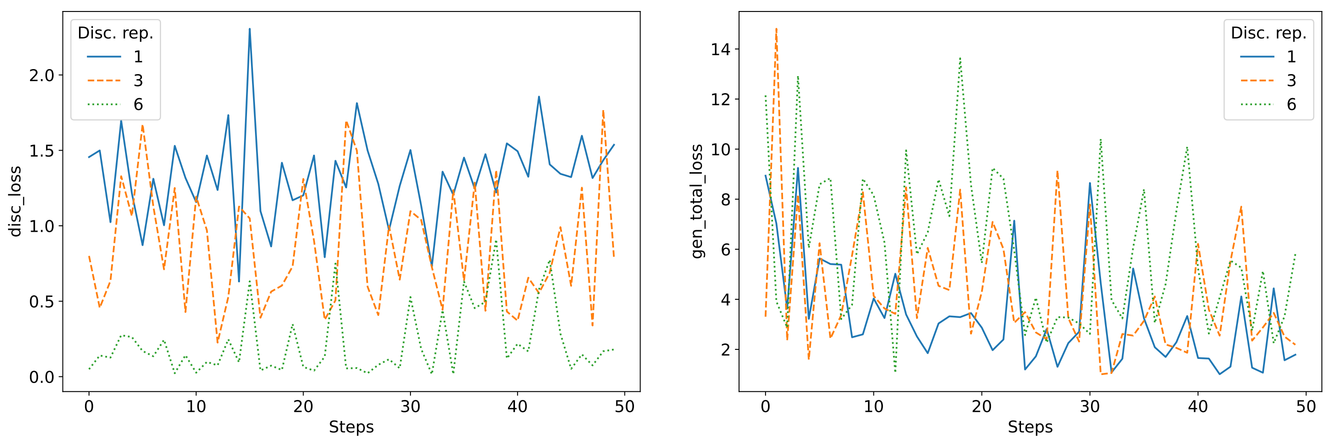

Figure 11. The number of episodes of Discriminator training per every episode of Generator training (termed Discriminator repetition or Disc. rep.), understandably, has a higher impact on the Discriminator loss than on the Generator loss. The left and the right panel show this effect on the Discriminator loss and the Generator loss respectively.

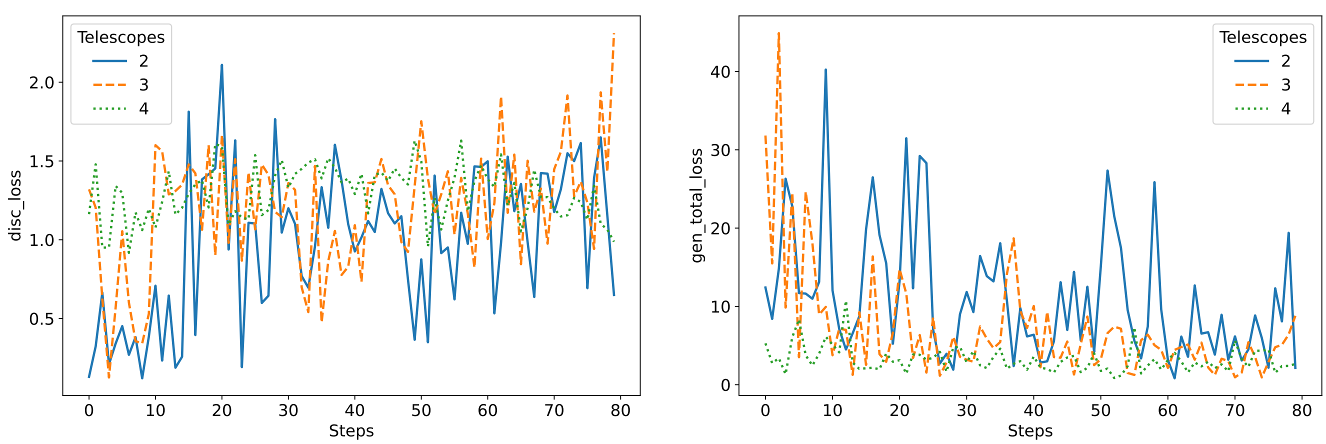

Figure 12. Discriminator and Generator loss for different numbers of telescopes. It has a very significant impact on the model performance. If there are only two telescopes, both Discriminator and Generator are not trained smoothly. The result of four telescopes is a lot better because the loss functions change only slightly with increasing steps.

Subsequently, the GAN framework was extended to a conditional model (M. Mirza et al. 2014). In this formulation, both the Generator and the Discriminator receive additional information y, and the value function of the conditional GAN (cGAN) is expressed as:

$$V(D,G) = \mathbb{E}_{x\sim p_{\mathrm{data}}(x)}[\log D(x|y)] + \mathbb{E}_{z\sim p_z(z)}[\log(1 - D(G(z|y)))] \tag{16}$$

P. Isola et al. (2017) further observed that combining the cGAN from Eq. (16) with the traditional L1 loss improves the results, as the Generator is encouraged to produce outputs closer to the ground truth. Hence, the function that is minimized is:

$$L_{tot} = \arg\min_G \max_D V(D,G) + \lambda \cdot L_1(G) \tag{17}$$

with the choice $\lambda = 100$ and

$$L_1(G) = \mathbb{E}_{x,y,z}[||y - G(x,z)||_1] \tag{18}$$

This type of network has demonstrated remarkable robustness across a variety of applications. For example, it can generate colored images from grayscale inputs based on architectural labels, transform images from day to night, and even predict maps from satellite data. A more extensive list of applications is provided in P. Isola et al. (2017).

3.1. *Generator*

As discussed above, in a GAN the Generator is responsible for producing synthetic data—in this case, images that resemble those of a fast-rotating star. In this work, the Generator is implemented as a U-Net convolutional network (O. Ronneberger et al. 2015). In such architectures, the image's spatial resolution is first reduced through downsampling and then restored via upsampling, resulting in a U-shaped structure. The downsampling process typically involves convolutional layers followed by a strided operation (with a stride of 2) to effectively subsample the image, and a leaky version of the Rectified Linear Unit (LeakyReLU) is employed as the activation function.

In contrast, the upsampling process uses only the standard Rectified Linear Unit (ReLU) for neuron activation. This stage also comprises convolutional layers followed by operations with a stride of 2 to upscale the image to a higher resolution. Additionally, a dropout layer is introduced at the beginning of the upsampling phase to mitigate overfitting of the Generator model (P. Isola et al. 2017).

After generating images, the Generator aims to deceive the Discriminator into classifying the generated images as real. The extent to which the Generator succeeds in this deception is quantified by the GAN loss. When the Discriminator is unable to distinguish between the generated and real images (i.e., when the GAN loss is minimized), the Generator is considered to have reached an optimal state. Conversely, if the generated image fails to fool the Discriminator, the Generator produces a new image for further comparison with the real image. Additionally, the Generator's performance is evaluated using another metric known as the L1 loss, which is defined as the mean absolute error between the pixels of the real image and those of the generated image. Balancing the minimization of both the GAN loss and the L1 loss enables the Generator to produce images that are not only realistic but also faithful to the input data.

3.2. *Discriminator*

The Discriminator is tasked with classifying the images produced by the Generator as either real or fake. It takes a real image from the dataset (often referred to as the target image for the Generator) and provides feedback to guide the Generator toward producing more accurate images. In this work, the PatchGAN model (P. Isola et al. 2017) is employed as the Discriminator. Unlike a traditional global classifier, PatchGAN evaluates individual patches of the image, outputting a grid of predictions rather than a single scalar value.

The Discriminator's architecture begins with an initializer that accepts both the input (generated) images and the corresponding real images. Initially, Salt-and-Pepper noise is added to the input images. PatchGAN then reduces the spatial dimensions of the images to extract localized features, ensuring the model focuses on smaller regions. In this downsampling stage, a leaky version of the Rectified Linear Unit (LeakyReLU) is applied in the convolutional layers, similar to the approach used in the Generator.

Subsequently, zero padding is applied—adding rows and columns of zeros around the images—to prevent the loss of spatial information during convolution and to facilitate the extraction of deeper features from the downsampled output. Following this, batch normalization is employed to stabilize learning by normalizing activations, and the Discriminator begins classifying each patch as real or fake. This is followed by additional layers involving LeakyReLU activation, zero padding, and convolution, culminating in a final prediction that the Generator uses as feedback.

The effectiveness of the Discriminator is measured by its ability to distinguish between real and fake images, quantified through the Discriminator loss. This loss is composed of two parts: one that measures how accurately the Discriminator identifies real images (by comparing predictions to a target value of 1) and another that assesses how accurately it identifies fake images (by comparing predictions to a target value of 0). Together, these loss components ensure that the Discriminator improves its classification performance, which in turn challenges the Generator to produce increasingly realistic images.

4. NETWORK PARAMETERS

Here, we discuss the parameters of the GAN architecture used for reconstructing images of stellar objects using II. Given the adversarial nature of GANs—where the Generator and Discriminator engage in a min-max game—careful tuning of key parameters is critical to ensure that both networks are well-balanced for effective training.

4.1. *Data Preparation*

First, we simulate fast-rotating stars, modelling them as oblate spheroids with varying radii and an oblateness ranging between 0.5 and 1. We also consider different viewing angles, assuming a linear dependence for the effect of gravity darkening. The traced ellipses result from integrating over the source's hour angle. For hyperparameter tuning and comparing different telescopes, the total observing duration is set to approximately 11.5 hours. Finally, the ellipses are plotted, converted into grayscale images, resized, and stored as raw arrays to facilitate further analysis.

Next, Salt and Pepper noise is introduced usually at a rate of 0.5%. Then, the images are resized and their mean is subtracted. A two-dimensional Fast Fourier Transform, along with a Fourier shift, is applied, yielding a complex number for each pixel. Since II does not measure phase, the absolute value is calculated (as shown in Fig. 4 on both linear and logarithmic scales for visualization).

Next, sparse sampling is introduced via pixel-wise multiplication between the absolute-valued Fourier-transformed image (Fig. 4) and the sparse sampling map (Fig. 2). The result is a map in the Fourier plane featuring several ellipses, which is also referred to as the sparse sampling map (Fig. 5). This map represents the sparse sampling of the signal space (Fig. 4) corresponding to the source (Fig. 3) observed with four telescopes (Fig. 1).

Finally, the pixels are normalized and converted to 8-bit integers. This image represents the sparsely sampled phaseless visibility as it can be measured with II. The image shown in Fig. 5 serves as the input for the GAN, which also requires the corresponding ground truth image. Consequently, the simulated stars are resized using the same algorithm and converted to 8-bit integers to reduce bias. The GAN must have access to the ground truth corresponding to each input image; therefore, the input and ground truth images are merged side-by-side (as shown in Fig. 6) and used to train the GAN. This procedure is applied to all simulated stars, with 10% used as test data, 10% as validation data, and the remaining 80% as training data.

4.2. *GAN Architecture*

The GAN used in this work is based on pix2pix, which utilizes a conditional GAN (cGAN) as discussed in the previous section (P. Isola et al. 2017). This architecture is highly robust and has already been applied to various problems. For instance, the TensorFlow tutorials[6] demonstrate its application to a dataset of architectural facades. However, to adapt the pix2pix GAN for the phase retrieval problem, some modifications are necessary. The network is implemented using the TensorFlow library (M. Abadi et al. 2016), calculations are performed with scipy (P. Virtanen et al. 2020), and plots are generated with matplotlib (J. D. Hunter 2007).

4.3. *Hyperparameter Tuning*

The GAN used in this work depends on several parameters, which are explained briefly below (for a more in-depth discussion, see K. P. Murphy 2022).

The learning rate of the optimizer determines how much the model updates its parameters with each iteration. A learning rate that is too small may lead to underfitting, while one that is too large can render the model unstable. Therefore, selecting an appropriate learning rate is crucial (K. P. Murphy 2022). Fig. 7 illustrates the effect of different learning rates on both the Generator and Discriminator losses. As expected, lower learning rates result in fewer outliers in the loss functions, indicating more stable updates. Although all models eventually stabilize at a similar level, lower learning rates are preferred.

The kernel size refers to the dimensions of the convolutional kernel used in the network, determining how many pixels are combined to produce a new pixel. A larger kernel size can capture features spanning several pixels, but it may also incorporate unrelated features. As shown in Fig. 8, the kernel size does not have a significant impact on the loss functions; however, smaller kernel sizes tend to produce more outliers, suggesting

[6] https://www.tensorflow.org/tutorials/generative/pix2pix

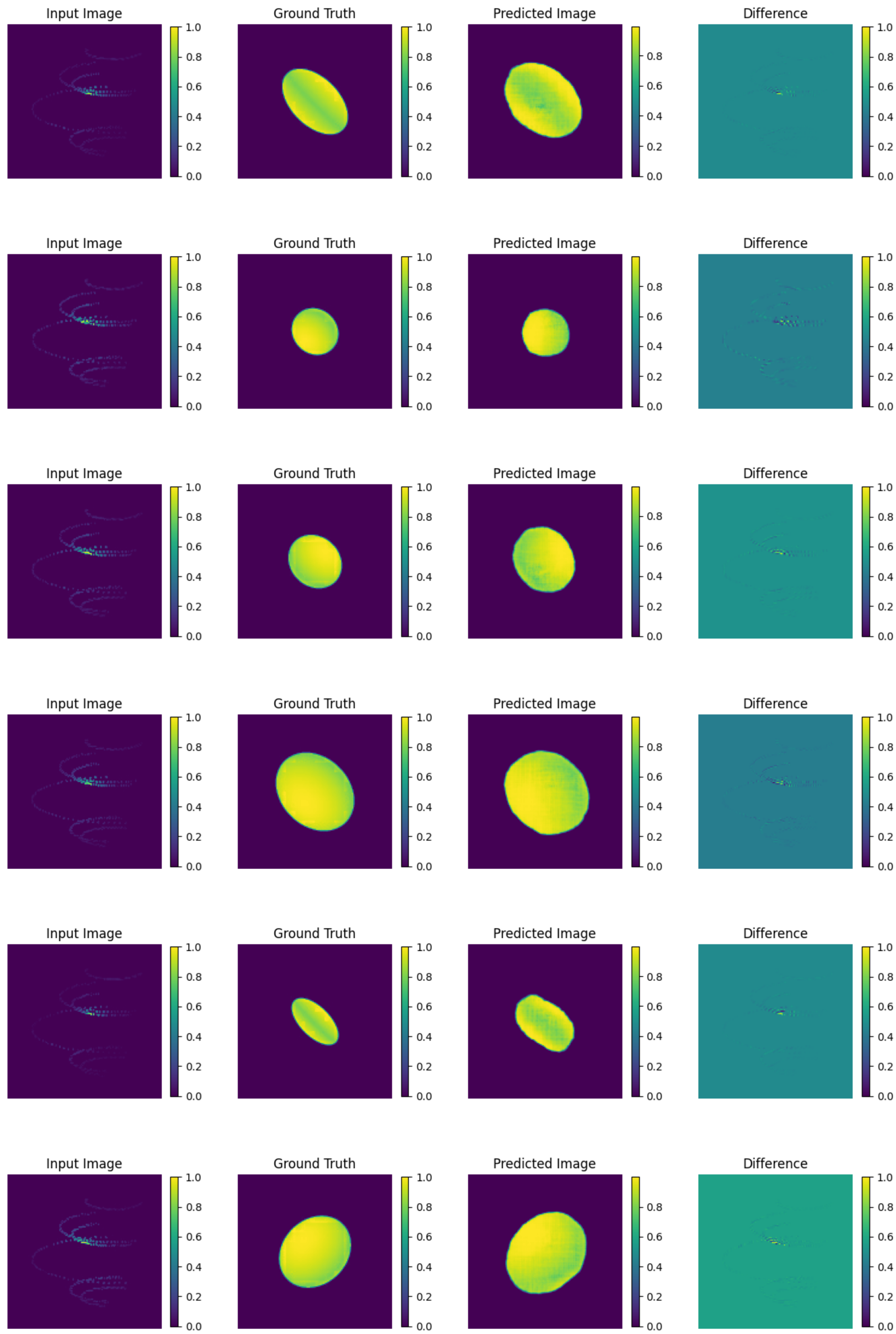

Figure 13. Example results of image reconstruction using the GAN model along with the II observations simulated in this work. Each row in this figure represents the results for a hypothetical fast-rotating star. Going from left to right in each row, the first panel represents the simulated (u, v) plane II signals obtained using the six baselines (of Fig.5). This image is presented, as the input, to the trained GAN to produce a predicted image. The second panel is the real image of the star, also called the ground truth. The third panel is the reconstructed image, or the predicted image, produced by the trained GAN model . The fourth and the last panel is the difference between the ground truth and the predicted image in the (u, v) plane. The uniform background color in the difference panel suggests that the network is picking up white noise from the (u, v) plane.

that either the Generator or Discriminator may gain an advantage. Therefore, a kernel size of 5 is preferred.

The amount of noise is controlled by two parameters, "alpha" and "beta", which indicate the percentage of pixels altered to either white or black — hence the term Salt and Pepper noise. Here, "alpha" is applied to the real image, while "beta" is applied to the generated image. Different ratios ("alpha/beta") can lead to varying model performance; however, our results indicate that distinct noise rates do not significantly affect the loss functions. Figure 9 shows the loss functions for smaller images (64×64), and due to the negligible impact, this analysis was not repeated for larger images.

The batch size defines the number of images processed simultaneously by the network. Smaller batch sizes have been observed to improve generalization (S. J. Prince 2023). As illustrated in Fig. 10, processing two images at once results in fewer outliers. However, because a larger batch size significantly increases training time, a batch size of 1 is used.

When training GANs, one strategy to potentially boost performance is to give the Discriminator an advantage by increasing its number of training steps before returning to the Generator's training. While this can lower the Discriminator loss—as shown in Fig. 11—it also increases training time and leads to a slight rise in the Generator loss. Since the generated images do not noticeably improve with additional Discriminator training, both networks are typically trained with the same number of steps.

Finally, the degree of sparse sampling can be varied to provide the model with access to more pixels. Increasing the number of telescopes results in more baselines and, consequently, more available pixels. Fig. 12 shows the loss functions for different numbers of telescopes. There is a significant disparity in performance, partly because the relationship between telescopes and baselines is not linear. For example, the Fourier plane can be sampled along six tracks when using four telescopes, as compared to only one track if using only two telescopes. In the case of two telescopes, both the Generator and Discriminator exhibit less smooth training, as indicated by the outliers. Performance improves with three telescopes and becomes very promising with four. Overall, the degree of sparse sampling appears to have the most pronounced effect of all the hyperparameters.

5. IMAGES RECONSTRUCTED BY THE GAN

In this section, we begin by discussing phase retrieval using hyperparameters, as mentioned earlier, followed by an analysis in which multiple sources are trained simultaneously. The best performance for image reconstruction has been observed with a learning rate of $2 \cdot 10^{-4}$, a kernel size of $5{\times}5$, and equal noise percentages (alpha/beta = 1) applied to both the original and generated images. A batch size of 1 is used, and equal training is provided to both the Discriminator and the Generator.

5.1. *Predicted Image from the Trained GAN*

Fig. 13 demonstrates the success of the GAN in training a model to reconstruct the images of fast-rotating stars using their Intensity Interferometry (II) observation . The GAN was trained on the training datasets for 60,000 steps and subsequently tested on various validation datasets to produce predicted images of the stars. The training was performed on a CPU using two nodes, each with 96 threads, and required approximately nine hours. In Fig. 13, four combined images illustrate the GAN's performance in reconstructing the stars' shape, size, and brightness distribution using II.

- The left panel shows the signals collected from six baselines, which serve as the input for the Generator during training.

- The first middle panel displays the real image, or ground truth, which the Disciminator loss function uses to distinguish from the images generated by the Generator. During training, the GAN aims to keep minimum difference between these ground truth image and generated images.

- The second middle panel presents the reconstructed, or predicted, image produced by the trained GAN, highlighting its success in image reconstruction.

- The right panel shows the difference between the ground truth and the predicted image in the interferometric plane. The uniform background color of the frames suggests that the GAN model picks up white noise in the interferometric plane.

The predicted images in Fig. 13 yield encouraging results, accurately conveying visual information about the source's size, shape, and brightness distribution across its surface using only six baselines. However, further improvements can be achieved by increasing the number of telescopes to maximize coverage of the (u, v) plane, making the existing and upcoming CTAO an ideal candidate for this approach.

5.2. *Evaluation of GAN using Moments*

The reconstructed images are visually compelling, demonstrating the GAN's effectiveness in using II to reconstruct images. However, visual assessment alone is

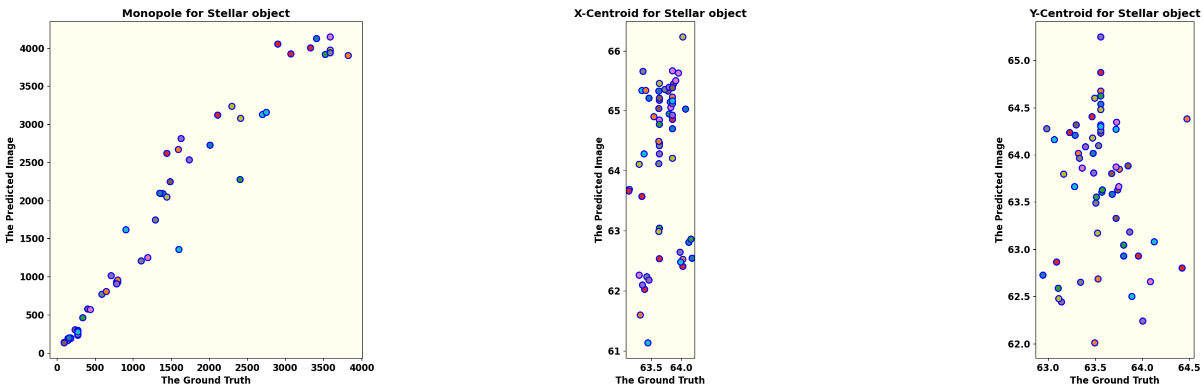

Figure 14. This set of figures shows the comparison of monopole, x-centroid, and y-centroid for ground truth and predicted images generated by trained GAN.

insufficient; statistical evaluation is necessary to validate the results. To achieve this, we employ image moments as a statistical method. Image moments capture key properties of the reconstructed objects—such as shape, size, and intensity distribution—by quantifying features like position, orientation, and brightness distribution. By comparing the moments of the GAN-generated images to those of the ground truth, we can objectively assess the consistency and accuracy of the reconstruction. This approach provides a reliable framework for evaluating reconstruction quality, as image moments can reveal subtle differences in geometric and intensity properties that might not be apparent through visual inspection alone.

The raw moment M_{ij} of an image is defined as (M.-K. Hu 1962)

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y) \tag{19}$$

where $I(x, y)$ represents the intensity at pixel (x, y). The zeroth order raw moment, or monopole, represents the total intensity of an image. It is computed by summing all pixel values across the image, yielding an overall intensity measure. In this context, analyzing the monopole provides the total flux of fast-rotating stars. According to Eq. (19), the monopole of an image is calculated as:

$$M_{00} = \sum_x \sum_y I(x, y). \tag{20}$$

The left figure of Fig. 14 displays the monopole values for 50 reconstructed images. The plot reveals a linear relationship between the monopole of the ground truth (real image) on the x-axis and that of the predicted (reconstructed) image on the y-axis, consistent across sources of varying shapes and sizes. This linearity confirms that the predicted images have an overall intensity (flux) that closely matches the ground truth. However, while the monopole effectively represents the total brightness, it does not provide information about the position, shape, size, or detailed brightness distribution of the fast-rotating stars. For these aspects, higher-order moments are necessary.

The center of mass of the sky image of any stellar object, is determined by its centroid, which provides the x and y coordinates representing the spatial position of the image. This centroid is computed using the first-order raw moments in conjunction with the monopole (the zeroth-order moment). The formulations for the centroid along the x and y directions are given by:

$$x_c = \frac{M_{10}}{M_{00}} = \frac{\sum_{x,y} x \cdot I(x, y)}{\sum_{x,y} I(x, y)}$$

$$y_c = \frac{M_{01}}{M_{00}} = \frac{\sum_{x,y} y \cdot I(x, y)}{\sum_{x,y} I(x, y)} \tag{21}$$

Here, $I(x, y)$ represents the intensity at pixel (x, y), M_{00} is the monopole (total intensity), and M_{10} and M_{01} are the first-order raw moments along the x and y axes respectively. This formulation accurately captures the spatial center of mass of the stellar object in the image.

The middle and right panel of Fig. 14 compare the centroids (x_c, y_c) of 50 predicted images with their corresponding ground truths respectively. The clustering of centroids within a specific scale range across all results indicates that the reconstructed images accurately represent the spatial location of the fast-rotating star relative to the ground truth.

Furthermore, these calculated centroids are instrumental in analyzing the shape, size, and brightness distribution of the stars using higher-order image moments. To this end, the central moment of an image is calculated according to:

$$\mu_{pq} = \frac{1}{M_{00}} \sum_x \sum_y (x - x_c)^p (y - y_c)^q I(x, y). \tag{22}$$

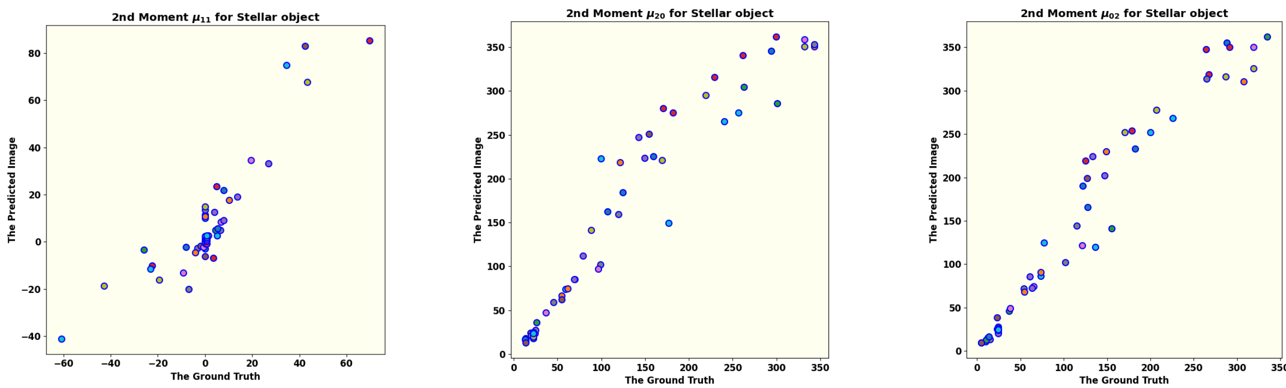

Figure 15. The second-order central moments provide information about the size and shape of stellar objects. Shown here are all the second-order central moments for ground truth and predicted images generated by the trained GAN. From left to right these are $\mu_{11}, \mu_{20}, \mu_{02}$.

The sum of p and q defines the order of the central moment.

Fig. 15 presents the second-order central moments ($\mu_{11}, \mu_{20}, \mu_{02}$), which are used to study the structure of a fast-rotating star along the line of sight (as explained in the upcoming subsection). All three plots demonstrate a linear relationship in the second-order moments, similar to the monopole, thereby confirming the success of applying the GAN to reconstruct images with II.

The brightness distribution is characterized by the skewness of the image, which is quantified by calculating the third-order central moments ($\mu_{30}, \mu_{03}, \mu_{21}, \mu_{12}$). Fig. 16 presents all third-order moments for both the ground truth and the reconstructed image. The skewness along the x and y axes (μ_{30} and μ_{03}) appears acceptable, as shown in both upper panel of Fig. 16, where a linear relationship exists between the ground truth and predicted images. However, the other higher-order moments (μ_{21} and μ_{12})—particularly μ_{12}, as depicted in both lower panel of Fig. 16—do not align as well. This indicates that further improvement is possible and should be investigated.

5.3. *The reconstructed Parameters for object*

The centroids (x_c, y_c) indicate only the center of the star and its spatial location in the image. In contrast, the second-order central moments determine the orientation, semi-major axis, and eccentricity relative to the source's center (M. R. Teague 1980). These moment-based parameters fully describe the two-dimensional ellipse that fits the image data.

The orientation of a fast-rotating star along the line of sight is defined in terms of second-order central moments as

$$\theta = \frac{1}{2}\arctan\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right). \tag{23}$$

The semi-major and semi-minor axes of the stellar object are computed using the second-order central moments and are denoted as a and b, respectively.

$$\begin{aligned} a &= 2\sqrt{mp + \delta} \\ b &= 2\sqrt{mp - \delta} \end{aligned} \tag{24}$$

where,

$$mp = \frac{\mu_{20} + \mu_{02}}{2} \tag{25}$$

and

$$\delta = \frac{\sqrt{4\mu_{11}^2 + (\mu_{20} - \mu_{02})^2}}{2}. \tag{26}$$

Using the calculated axis values, the eccentricity of the fast-rotating star is determined as:

$$e = \sqrt{1 - a/b}. \tag{27}$$

Eqs. 23-27 describe the elliptical nature of the stellar object (in this case, a fast-rotating star) and provide information on its shape and size, depending on the computed values. In contrast, the brightness distribution is characterized by skewness, which is quantified using third and higher-order moments.

6. CONCLUSION

Intensity Interferometry (II) is re-emerging as a promising technique to overcome the challenges of very long baseline interferometry in the optical wavelength range. However, compared to radio-interferometry, optical interferometry faces an important hurdle: photon correlation captures only the magnitude of the interferometric signal, resulting in a loss of phase information.

This work addresses the challenge of phase retrieval in II using a machine-learning technique, specifically a conditional Generative Adversarial Network (cGAN).

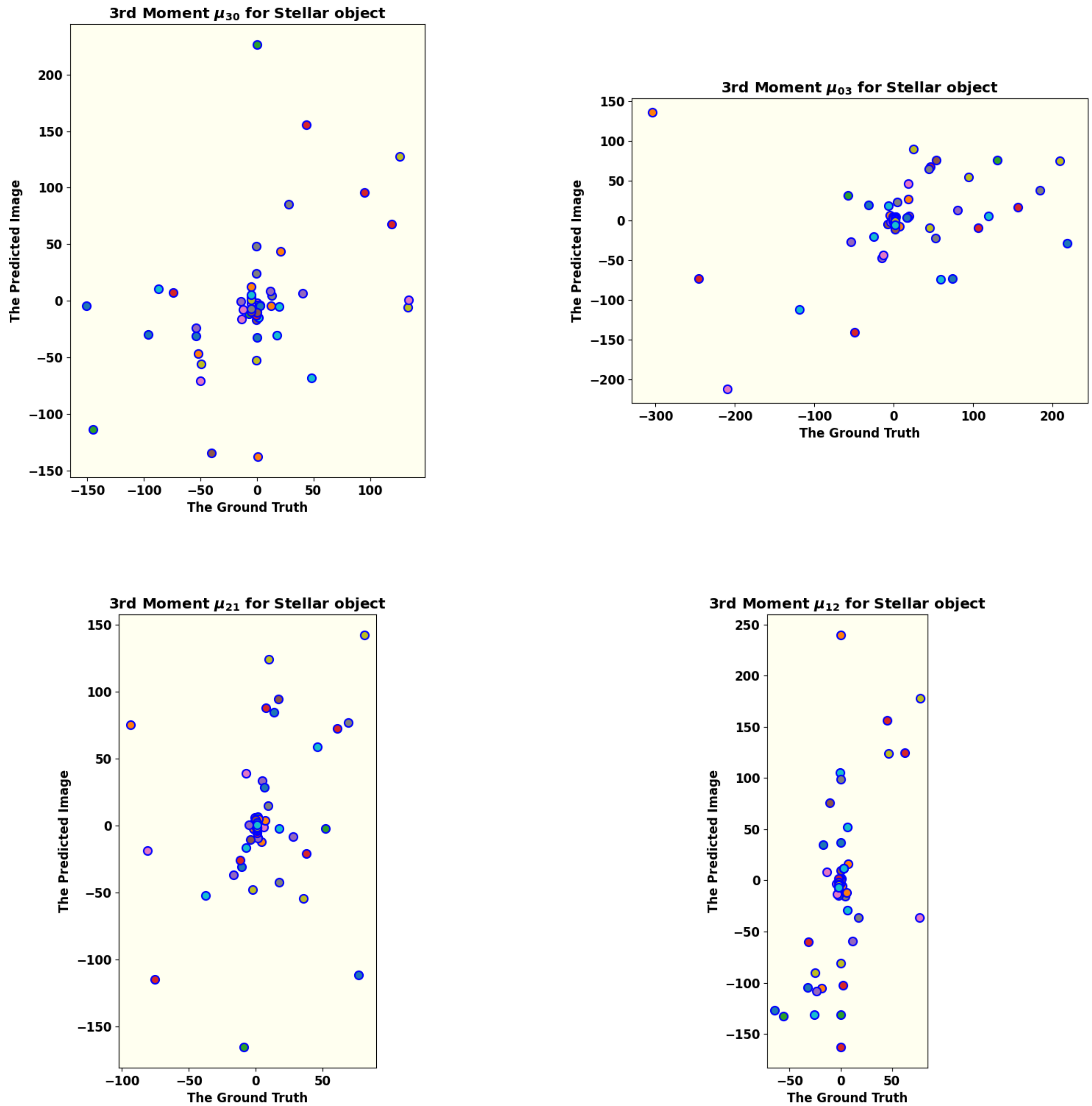

Figure 16. Shown here are all the third-order central moments for ground truth and predicted images generated by the trained GAN. They represent the skewness of the brightness distributions. The panels in reading order show $\mu_{30}, \mu_{03}, \mu_{21}, \mu_{12}$.

Our study demonstrates that applying a cGAN to II data successfully recovers the size, shape, and brightness distribution of a fast-rotating star. Evaluations based on image moments—specifically, the monopole, second, and third-order moments—support the effectiveness of cGAN in achieving accurate image reconstruction from a simulation of II from a single site with four telescopes.

While the results of this study highlight the significant potential of machine learning, and in particular the applicability of cGAN, for image reconstruction in II, several aspects require further refinement. First, an important factor in the reconstruction process is the extent of Fourier plane coverage, which depends on the number of available telescopes and the total observing time. The reasonable success of this piece of work suggests that a network of higher number of telescopes providing higher number of baselines and greater coverage of the (u, v) plane signal, projects of image reconstruction of more complicated stellar systems can be undertaken. Future work might explore different observatory layouts to assess their impact on image reconstruction quality. Second, detector efficiencies, which impact the signal-to-noise ratio (SNR) of actual observational data, have not yet been incorporated; addressing these factors will be crucial for more accurate SNR estimation. Third, exploring and comparing alternative methods for image generation could reveal approaches that outperform cGAN in reconstructing stellar images with II. Fourth, experimenting with different loss functions could provide additional insights into the reconstruction quality. Although further testing is needed to refine the GAN and enhance its robustness and reliability, our findings suggest that machine learning is a promising approach for phase reconstruction in II.

ACKNOWLEDGEMENTS

Acknowledgment: One of the authors (SS) gratefully acknowledges the computing facilities and the local hospitality extended to him by the Inter-University Centre for Astronomy and Astrophysics (IUCAA), Pune, India under its Visiting Associate Programme during the preparation of this manuscript.

REFERENCES

Abadi et al., M. 2016, arXiv preprint arXiv:1603.04467

Abe et al., S. 2024, MNRAS, 529, 4387

Acciari et al., V. A. 2020, MNRAS, 491, 1540

Acharyya et al., A. 2024, The Astrophysical Journal, 966, 28

Aleksić et al., J. 2016, Astroparticle Physics, 72, 76

Archer, A., Aufdenberg, J. P., Bangale, P., et al. 2025, arXiv e-prints, arXiv:2506.15027, doi: 10.48550/arXiv.2506.15027

Baumgartner et al., S. 2020, MNRAS, 498, 4577

Coccomini et al., D. 2021, arXiv preprint arXiv:2122.11578

Dravins et al., D. 2013, Astroparticle Physics, 43, 331, doi: 10.1016/j.astropartphys.2012.04.017

Fienup, J. 1982, Applied Optics, 21, 2758

Gamo, H. 1963, Journal of Applied Physics, 34, 875

Gerchberg, R. W. 1972, Optik, 35, 237

Glauber, R. J. 1963, Physical Review, 130, 2529

Goldberger et al., M. L. 1963, Physical Review, 132, 2764

Goodfellow et al., I. 2014, Advances in neural information processing systems, 27

Hanbury Brown, R., & Twiss, R. Q. 1956, Nature, 178, 1046

Hanbury Brown et al., R. 1954, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 45, 663

Hanbury Brown et al., R. 1957, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 242, 300

Hanbury Brown et al., R. 1958, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 243, 291

Hanbury Brown et al., R. 1974, Monthly Notices of the Royal Astronomical Society, 167, 121

Hecht, E. 2002, Optics (Addison Wesley)

Holmes et al., R. 2010, in Adaptive Coded Aperture Imaging, Non-Imaging, and Unconventional Imaging Sensor Systems II, Vol. 7818, SPIE, 175–185

Hu, M.-K. 1962, IRE transactions on information theory, 8, 179

Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90, doi: 10.1109/MCSE.2007.55

Isola et al., P. 2017, in Proceedings of the IEEE conference on computer vision and pattern recognition, 1125–1134

Kirisits et al., C. 2024, arXiv preprint arXiv:2406.14143v2

Le Bohec, S; Holder, J. 2006, The Astrophysical Journal, 649, 399

Leonard et al., M. 1995, Optical Coherence and Quantum Optics (Cambridge University Press)

Li et al., X. 2014, in Imaging Spectroscopy, Telescopes and Large Optics, Vol. 9298, SPIE, 92981G1–G7

Lucy, L. B. 1967, Zeitschrift für Astrophysik, Vol. 65, p. 89, 65, 89

Maeder, A. 1999, Astronomy & Astrophysics, 347, 185

McAlister et al., H. A. 2005, The Astrophysical Journal, 628, 439

Mirza et al., M. 2014, arXiv preprint arXiv:1411.1784

Murphy, K. P. 2022, Probabilistic machine learning: an introduction (MIT press)

Mustafa et al., M. 2019, Computational Astrophysics and Cosmology, 6, 1

Nuñez et al., P. D. 2010, in Optical and Infrared Interferometry II, Vol. 7734, SPIE, 458–467

Nuñez et al., P. D. 2012a, Monthly Notices of the Royal Astronomical Society, 419, 172

Nuñez et al., P. D. 2012b, Monthly Notices of the Royal Astronomical Society, 424, 1006

Prince, S. J. 2023, Understanding deep learning (MIT press)

Rai et al., K. N. 2021, Monthly Notices of the Royal Astronomical Society, 507, 2813, doi: 10.1093/mnras/stab2391

Rai et al., K. N. 2022, Monthly Notices of the Royal Astronomical Society, 516, 2864, doi: 10.1093/mnras/stac2433

Ronneberger et al., O. 2015, in Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18, Springer, 234–241

Sato et al., T. 1978, Applied optics, 17, 2047

Sato et al., T. 1979, Applied Optics, 18, 485

Sato et al., T. 1981, Applied Optics, 20, 2055

Schawinski et al., K. 2017, MNRAS, 467, L110

Teague, M. R. 1980, Josa, 70, 920

Teague, M. R. 1983, Journal of Opticala Society of America, 73, 1434

Virtanen et al., P. 2020, Nature methods, 17, 261

Vogel et al., N. 2025, MNRAS, 537, 2334

Von Zeipel, H. 1924, Monthly Notices of the Royal Astronomical Society, Vol. 84, p. 665-683, 84, 665

Zhang et al., J. 2020, Opt. Lett., 45, 3649