

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

2 KM NITU RAI,<sup>1,2</sup> YURI VAN DER BURG,<sup>3</sup> SOUMEN BASAK,<sup>1</sup> PRASENJIT SAHA,<sup>3</sup> AND SUBRATA SARANGI<sup>4,5</sup>

3 <sup>1</sup>*School of Physics, Indian Institute of Science Education and Research Thiruvananthapuram, Maruthamala PO, Vithura,*  
4 *Thiruvananthapuram 695551, Kerala, India.*

5 <sup>2</sup>*Aryabhatta Research Institute of Observational Sciences, Manora Peak, Nainital 263129, India.*

6 <sup>3</sup>*Physik-Institut, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland.*

7 <sup>4</sup>*School of Applied Sciences, Centurion University of Technology and Management, Odisha-752050, India.*

8 <sup>5</sup>*Visiting Associate, Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, Pune 411 007, Maharashtra,*  
9 *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 II. This method successfully reconstructs the shape, size, and brightness distribution of simulated, fast-rotating stars based on a sparsely sampled spatial power spectrum obtained from a hypothetical ground-based II facility composed of four Imaging Atmospheric Cherenkov Telescopes (IACTs). 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

Humans instinctively feel a relationship with the stars. One of the primary scientific projects of humanity is to figure out what the stars are and how do they do what they do. The first obvious step in this project, beyond measuring their global parameters like diameter, mass, orbital and astrometric elements, is to obtain images of the stars with all details of the stellar surfaces. In case of the Sun, this is done routinely by observatories and Sun-observation satellites. But such a routine still remains a challenge even for the  $\alpha$ -Centauri system, our nearest stellar neighbour. The objective is to achieve capability of high fidelity image reconstruction of distant stars. Two interferometry based techniques, namely, Michelson Interferometry (MI) or Intensity Interferometry (II) have emerged during the last century to address this objective. A discussion of the development of these two approaches and comparison of their respective merits and challenges can be found in (K. N. Rai et al. 2025). The work reported here presents the results of a first effort at applying a Conditional Generative Adversarial (neural) Network (cGAN) to image reconstruction of a fast rotator using its simulated II observations.

The foundational basis of II stems from the pioneering experiments and theoretical investigations carried out initially by Hanbury Brown and Twiss (HBT) (R. Han-

bury Brown & R. Q. Twiss 1956; R. Hanbury Brown et al. 1957, 1958) and, later, by (R. J. Glauber 1963) and others. The correlation between photons (called the , widely referred to as the “HBT effect”) measured by a pair of photon detectors in two partially coherent beams of light was reported in 1956 R. Hanbury Brown & R. Q. Twiss (1956). K. N. Rai et al. (2025) present a recent variant of this experiment, carried out with pseudo-thermal light. The HBT Effect and the related theoretical investigations laid the foundation for the modern field of Quantum Optics.

Hanbury Brown and his collaborators led the creation and installation of the historic II facility at Narrabri, Australia, and reported the measurement of angular diameters of 32 stars and a few multiple star-systems (R. Hanbury Brown et al. 1974). Soon after this work, however, II observations of stars was stalled for over four decades due to the limits of the then available photon detectors and data processing equipment. With gradual mitigation of such issues, 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) as a secondary science application of these facilities during moonlit nights. SII observations at VERITAS, MAGIC, and H.E.S.S. are now being re-

ported (e.g., A. Acharyya et al. 2024; S. Abe et al. 2024; N. Vogel et al. 2025). This approach has the potential to enhance the scientific output of existing IACT facilities, and especially of the upcoming Cherenkov Telescope Array Observatory (CTAO). Simulations (K. N. Rai et al. 2021, 2022, e.g.,) have shown that recent advancements in photon detectors could be effective in achieving high-precision measurements of parameters for stellar objects.

The thrust of these efforts has been to measure the average global parameters of stars and star systems. High-fidelity imagery, however, would transcend the measurement of global and average stellar parameters, such as angular diameters, binary separations, and orbital characteristics, which offer only an integrated view of the star or star system as a whole. Such imaging capabilities promise direct insights into the dynamic surface phenomena, including limb darkening, convection cells, granulation, star spots, oblateness and gravity darkening in rapid rotators and atmospheric structures, akin to the detailed observations routinely conducted on our own Sun.

As it stands today, studies grappling various issues of image reconstruction are being reported (X. Haubois et al. 2009; R. P. Norris et al. 2021; L.-C. Liu et al. 2024; L.-C. Liu et al. 2025). MI-based image reconstruction has made substantial progress in this area with efforts at generating constructed images of stars like Betelgeuse (X. Haubois et al. 2009) and AZ Cyg (R. P. Norris et al. 2021). On the other hand, II-based methods are at a nascent stage. Some recent publications L.-C. Liu et al. (2024); L.-C. Liu et al. (2025) have demonstrated, through outdoor experiments, imaging millimeter-scale targets at 1.36 km with a resolution 14 times better than a single telescope’s diffraction limit. A “flexible computational algorithm” reconstructs images from intensity correlations, overcoming atmospheric turbulence and optical imperfections.

We report here, the results of our attempt – the first of its kind – to construct the gravity-darkened sky-image of fast rotating stars consistent with their respective simulated ground-based II-observations using a cGAN neural network architecture. Image reconstruction in gravity-darkened fast-rotating stars has long been examined using various methods in MI (G. T. van Belle et al. 2001; A. Domiciano de Souza et al. 2003, 2005; H. A. McAlister et al. 2005; J. D. Monnier et al. 2007; E. Pedretti et al. 2009; M. Zhao et al. 2009; A. O. Martinez et al. 2021). Recently photosphere oblateness of  $\gamma$ -Cassiopeia (A. Archer et al. 2025) has been measured at the VERITAS observatory using II. These results put our work in context, and our work presented here is a natural next step especially of the work by A. Archer et al. (2025).

We implement a cGAN model (P. Isola et al. 2017) to reconstruct images of fast-rotating stars using their simulated Intensity Interferograms and sky-intensity distributions as input data for training, testing, and validation. We consider an array of four Imaging Cherenkov Telescopes (IACTs) whose relative positions approximately mimic those at VERITAS and simulate observations of a set of synthetically generated fast-rotating stars. The image predicted by the trained cGAN shows promising results in reconstructing the stars’ shapes and sizes. The reconstructed brightness distributions are then assessed using moments.

This paper is organized as follows. The next section discusses briefly the past efforts at image reconstruction on II, followed by the section on a discussion of II, focusing on its signal and noise characteristics for fast-rotating stars along the Earth’s rotation. The following section introduces the cGAN formulation and its structure. The fifth section details the parameter selection for training the network for image reconstruction. The sixth section presents the results of the trained network both visually and via image moments. Finally, the paper concludes with a discussion of the overall results.

## 2. PAST EFFORTS AT IMAGE RECONSTRUCTION RELEVANT TO INTENSITY INTERFEROMETRY

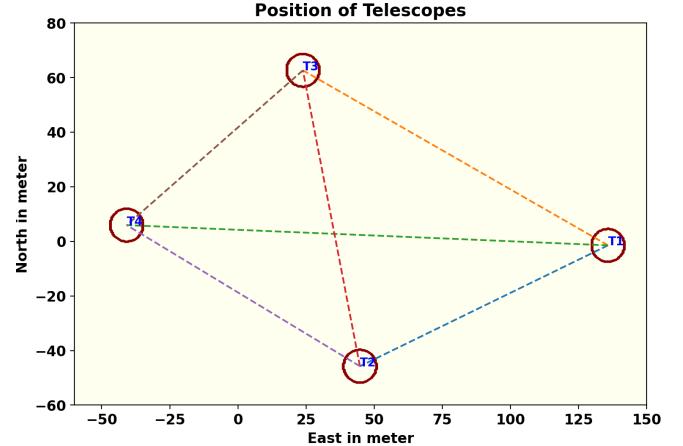
Several approaches have been developed for phase reconstruction in intensity interferometry (II). H. Gamo (1963) introduced triple-intensity correlation, applied by M. L. Goldberger et al. (1963) to microscopic systems and extended by T. Sato et al. (1978, 1979, 1981) to measure stellar diameters and phases, though limited by low signal-to-noise ratio (SNR). R. W. Gerchberg (1972) proposed an iterative phase retrieval method using image and diffraction plane data, sensitive to initial estimates and convergence speed. J. Fienup (1982) improved this with the Hybrid Input-Output algorithm, enhancing robustness in noisy conditions. R. Holmes et al. (2010) utilized Cauchy-Riemann relations for 1-D and 2-D image reconstruction, applied to simulated stellar data with Imaging Cherenkov Telescope Arrays (P. D. Nuñez et al. 2010, 2012a,b), but faced computational complexity for higher dimensions. X. Li et al. (2014) introduced a regularized iterative method incorporating priors (e.g., sparsity) to mitigate noise and ill-posedness, though challenged by parameter tuning and initial guess sensitivity. The Transport-of-Intensity Equation (TIE), proposed by M. R. Teague (1983), retrieves phase from intensity variations across planes; J. Zhang et al. (2020) solved TIE as a Poisson equation using a maximum intensity assumption, while C.

175 Kirisits et al. (2024) combined TIE with the Transport of Phase Equation for improved accuracy across  
 176 arbitrary apertures and non-uniform illumination, with convergence dependent on initial guesses and boundary  
 177 conditions.  
 178

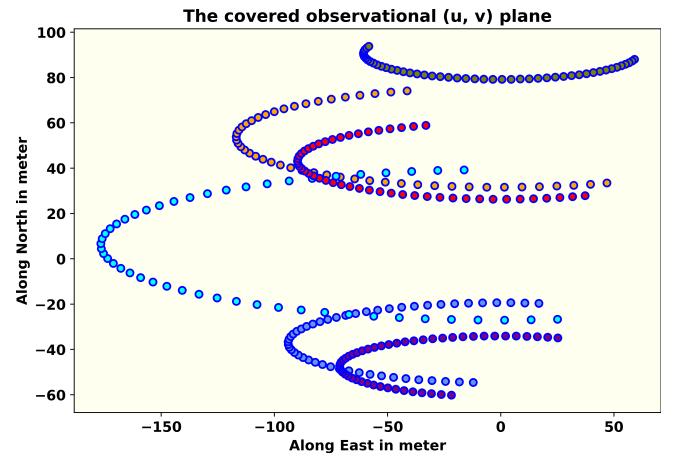
179 With non-linearity built into their architecture, artificial  
 180 neural networks (ANNs) empowered by deep learning  
 181 methods are promising for exploring the task of re-  
 182 constructing images of stellar objects from ground-based  
 183 observations. Convolutional Neural Networks (CNNs),  
 184 with their specialized architecture for processing two-  
 185 dimensional datasets, are a natural choice for image  
 186 processing tasks. In astronomical image reconstruction  
 187 projects, a common challenge is that the interferometric  
 188 data are typically undersampled as well as noisy. There-  
 189 fore, the CNN architectures and deep learning methods  
 190 employed must be capable of reliably learning both the  
 191 global context of the training dataset and the local fea-  
 192 tures within it. Among the various CNN architectures,  
 193 U-Net models (O. Ronneberger et al. 2015) have proven  
 194 successful in such tasks.  
 195

196 Furthermore, given that achieving a high signal-to-  
 197 noise ratio (SNR) is often challenging in astronomical  
 198 datasets, it is immensely beneficial if additional data can  
 199 be generated using the available information from the  
 200 observed sky density distribution and ground-based ob-  
 201 servations (II data, in our case) of the sources under in-  
 202 vestigation. Generative Adversarial Networks (GANs),  
 203 introduced by I. Goodfellow et al. (2014), have been  
 204 successful in such data augmentation tasks. Conditional  
 205 GAN (cGAN) architectures, proposed by M. Mirza et al.  
 206 (2014) and applied to a wide variety of datasets by P.  
 207 Isola et al. (2017), leverage additional information about  
 208 the images in the training datasets and have demon-  
 209 strated remarkable robustness in image recovery across  
 210 diverse data types.

211 In the astrophysical context, K. Schawinski et al.  
 212 (2017) employed a GAN model to recover features, such  
 213 as spiral arms, central bulges, and disk structures of  
 214 galaxies, from noise-affected images. M. Mustafa et  
 215 al. (2019) developed and customized a Deep Convo-  
 216 lutional GAN, dubbed “CosmoGAN”, capable of gen-  
 217 erating high-fidelity weak-lensing convergence maps of  
 218 dark matter distribution that statistically reproduce real  
 219 weak lensing structures. D. Coccioni et al. (2021)  
 220 have successfully generated credible images of planets,  
 221 nebulae, and galaxies using “lightweight” and “physics-  
 222 uninformed” GANs to produce synthetic images of cele-  
 223 stial bodies. They also generated a “Hubble Deep Field-  
 224 inspired” wide-view simulation of the universe.



**Figure 1.** A schematic representation of a hypothetical observation facility with an array of four Cherenkov Telescopes (IACTs) whose relative positions approximately mimic those at VERITAS. This array is used in simulating the II observation of the fast rotators. Each of the telescopes has a diameter of 12m. The baselines provided by the array are of the order of 100m.



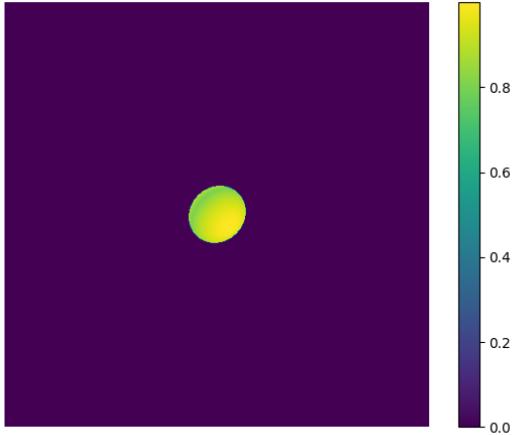
**Figure 2.** The tracks of the baselines due to Earth rotation (described in sec. 3.3) provided by the hypothetical observation facility of four telescopes arranged in fig. 1 for one night of observation. The number of baselines scales as square of number of telescopes in the array thus leading to greater coverage of the observational plane and better image reconstruction prospects.

### 225 3. INTENSITY INTERFEROMETRY (II) WITH 226 IACT ARRAYS

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

#### 231 3.1. *The signal for II*

232 As a simple example, let us consider a pair of IACTs  
 233 pointed at a star. Suppose the two telescopes simultane-



**Figure 3.** This figure shows one of the simulated fast rotating stars (FRS). The brightness is highest at the poles; gravity darkening visible along the equator. A total of 31460 such images of FRS with different parameter values have been generated to train the model.

ously measure the intensity of radiation  $I_1(t)$  and  $I_2(t)$ , respectively. The signals from these detectors are cross-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)

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

where  $\tau$  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).

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

where  $\Delta f$  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  $\Delta \nu \sim 1$  THz and  $\Delta f \sim 1$  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. It 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:

$$c = g^{(2)} - 1 = \frac{\Delta f}{\Delta \nu} |V_{12}|^2 \quad (3)$$

with  $|V_{12}|^2$  being a function of baseline  $b = \sqrt{u^2 + v^2}$  on the observational plane  $(u, v)$ . This implies the strength

of the signal would be enhanced if a larger number of baselines or pairs of telescopes are employed.

### 3.2. 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 received from a sky source onto their respective set of photo-multiplier tubes (PMTs, see e.g., J. Aleksić et al. 2016) with appropriate specifications needed for II observations. 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  $\lambda$  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

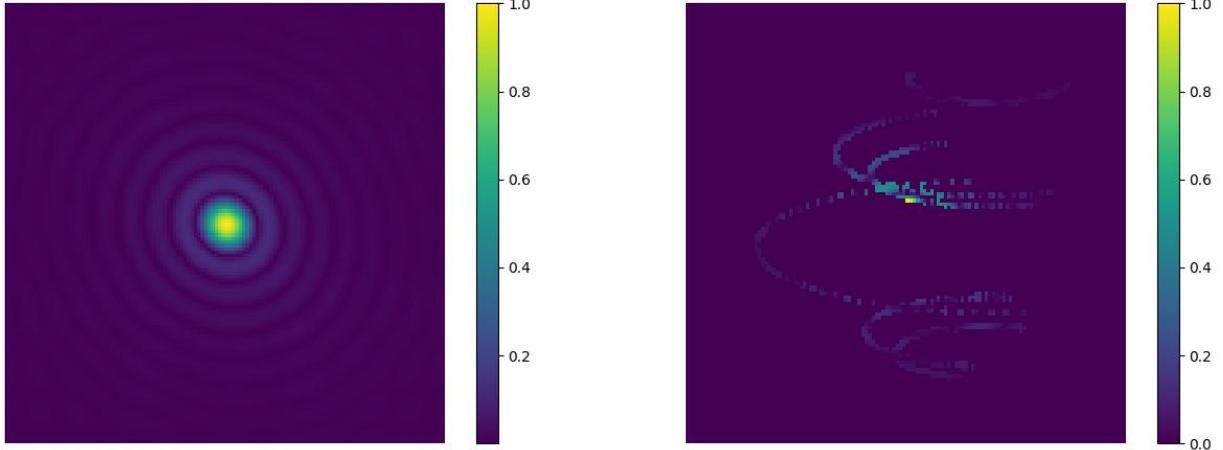
$$SNR = A \cdot \alpha \cdot q \cdot n \cdot F^{-1} \cdot \sigma \cdot \sqrt{\frac{T \Delta f}{2}} \cdot |V_{12}|^2 \quad (4)$$

Here,  $A$  is the total mirror area,  $\alpha$  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  $\sigma$  is the normalized spectral distribution of the light (including filters) (e.g., V. A. Acciari et al. 2020).

### 3.3. Baseline considerations

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

$$|V_{12}(b)|^2 = \frac{c(b)}{c(0)} \quad (5)$$



**Figure 4.** The left panel shows the absolute value of the two-dimensional Fast Fourier Transform of the source depicted in Fig. 3. It represents the intensity interferometric  $(u, v)$  plane image of the source that would be obtained by an infinite number of baselines (or an infinite number of telescopes observing the source). The right panel shows the absolute value of the same source measured along the tracks shown in Fig. 2. This figure reflects the sparse nature of the signal received by a realistic finite number of telescopes and baselines sampled from the full  $(u, v)$  plane signal space, as shown in the figure on the left panel. Both figures are plotted on a linear scale and normalized to the maximum pixel value in each respective figure.

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:

$$(u, v) = \frac{1}{\lambda} (b_E, b_N) \quad (6)$$

where  $b_E$  and  $b_N$  are, respectively, the baselines expressed in east and north coordinates. However, the 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. In order to gather maximum possible information on the source during the observation session and to cover as much of the observational plane as possible during such sessions, Earth's rotation must be taken into account using rotated baselines. For a given stellar source with declination  $\delta$  and hour-angle  $h$ , as observed by telescopes located at latitude  $l$ , equation (7) provides the rotated baselines for a given pair of telescopes (see e.g., eqs. 8–10 from S. Baumgartner et al. 2020) with the  $R$ -matrices representing the respective rotation operations.

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = R_x(\delta) \cdot R_y(h) \cdot R_x(-l) \begin{pmatrix} b_E \\ b_N \\ b_A \end{pmatrix} \quad (7)$$

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

$$\mathcal{N} = \frac{1}{2} N_T \cdot (N_T - 1) \quad (8)$$

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. Drawins et al. (2013) considered the telescope configurations then being planned and showed how it would provide a dense coverage of the interference plane.

#### 3.4. A Fictitious Fast Rotating Star: Our Test Case

In our work presented here, we attempt image reconstruction of a fast-rotating star using its simulated Intensity Interferometric observation in a cGAN architecture. Fast-rotating stars are important test cases for understanding various astrophysical processes, including stellar evolution, internal structure, and dynamical behaviour over time. 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 (and, therefore, the effective surface temperature) distributed across its surface. The brightness (effective temperature) is highest at the poles and lowest at the equator, a phenomenon known as gravity darkening (L. B. Lucy 1967). First direct interferometric detection of stellar photospheric oblateness (of Altair) was pioneered by G. T. van Belle et al. (2001) using the Palomar Testbed Interferometer (PTI) and the Navy Prototype Optical Interferometer (NPOI). Gravity darkening due to fast rotation was first observed through interferometric and spectroscopic data from the

CHARA Array for the fast-rotating star Regulus (H. A. McAlister et al. 2005). As pointed out earlier, these two pieces of work, all using Michelson Interferometry, make a subset of several others (G. T. van Belle et al. 2001; A. Domiciano de Souza et al. 2003; H. A. McAlister et al. 2005; A. Domiciano de Souza et al. 2005; J. D. Monnier et al. 2007; E. Pedretti et al. 2009; M. Zhao et al. 2009; A. O. Martinez et al. 2021). The first observation of photospheric oblateness (of  $\gamma$  Cassiopeiae or  $\gamma$ -Cas) using Intensity Interferometry (II) has been recently carried out at VERITAS observatory and is reported by A. Archer et al. (2025). Reportage of such observations of other  $\gamma$ -Cas like targets and other class of FRS by Cherenkov Telescope arrays, such as the MAGIC array, are expected by 2026. In addition, observation and measurement of gravity darkening using II is the natural next step and is yet to be reported. In this context, our work of reconstructing the image of FRSs from their II-simulated observations using cGAN is an attempt at solving this inverse problem along with mitigation of loss of phase information in II.

II 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. (5), as explained in subsection 3.1. The left panel of fig. 4 shows the signal from the source shown in Fig. 3 using II, displayed on linear scales. A 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 in right panel of Fig. 4), 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 a hypothetical observation facility having an array of four IACTs with their relative positions approximating those at VERITAS (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.

#### 4. GENERATIVE ADVERSARIAL NETWORKS

Generative Adversarial Networks (GANs) were introduced by I. Goodfellow et al. (2014). A GAN model involves two competing deep neural network models, referred to as the Generator and the Discriminator. These two networks engage in a zero-sum “Minimax” game, as in Game Theory. Given a real data set  $\{x_i\}$  (for example, a set of real images) drawn from some unknown distribution  $P_{\text{data}}(x)$  generated by some unknown or ill-

understood process, the objective here is to generate a new set (of images) whose probability distribution should match  $P_{\text{data}}(x)$  as closely as possible. During the training of the two models, the Generator samples a latent variable  $z$  from a known prior distribution  $P_z(z)$  (*e.g.*, the Normal Distribution) and produces a synthetic sample  $G(z)$  to start with and, subsequently, based on updates received from the Discriminator as its training progresses. The Discriminator, being a probabilistic binary classifier, receives either a real data sample  $x$  or a generated sample  $G(z)$ , and outputs a probability that the input is real. The Discriminator aims to maximise its classification accuracy, while the Generator aims to fool it by trying to minimize it. The training of these two networks proceeds alternately leading to the optimization of the adversarial loss function  $L(D, G)$  given by

$$\begin{aligned} L(D, G) &= \min_G \max_D V(D, G) \\ &= \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] \\ &\quad + \mathbb{E}_{z \sim p_z(z)} [\log (1 - D(G(z)))] . \end{aligned} \quad (9)$$

The two neural networks  $G(z) \equiv G(z; \theta_G)$  and  $D(x) \equiv D(x; \theta_D)$  are parameterized by  $\theta_i$  with  $i = G$  or  $D$  respectively. During its training, the Generator generates the differentiable function  $x_{\text{gen}} \equiv G(z)$  that maps  $z$  to the data space  $x$ . Through such maps the Generator generates its own distribution  $p_G(x_{\text{gen}})$  and through the training episodes, specifically by iteratively updating its parameters  $\theta_G$ , aligns this distribution with the distribution of real data  $p_{\text{data}}(x)$ . The training data set provided to the Discriminator is constructed by mixing real data points  $x$  and generated data points  $x_{\text{gen}}$  in equal ratio. The Discriminator generates the function  $D(x)$  that represents the probability that  $x$  originates from real data. Eq.(9) implies that training of the GAN model moves towards maximization of the expectation of  $D(x)$ . Through this process, both the parameters  $\theta_G$  and  $\theta_D$  are optimized such that  $p_G(x_{\text{gen}})$  gets maximally aligned with  $p_{\text{data}}(x)$ .

For a given fixed Generator  $G(z)$ , the problem can be reformulated as:

$$\begin{aligned} \max_D V(D, G) &= \mathbb{E}_{x \sim p_{\text{data}}} [\log D_G^*(x)] \\ &\quad + \mathbb{E}_{x \sim p_G} [\log (1 - D_G^*(x))] \end{aligned} \quad (10)$$

where  $D_G^*$  denotes the optimum of the discriminator. As seen in equation (11), the global optimum of equation (10) is achieved if and only if  $p_G = p_{\text{data}}$ . Furthermore, if both  $G$  and  $D$  are allowed to reach their respective optima, – the so called Nash point of the Minimax game –  $p_G$  converges to  $p_{\text{data}}$ .

$$D_G^*(x) = \frac{p_{\text{data}}(x)}{p_{\text{data}}(x) + p_G(x_{\text{gen}})} \quad (11)$$



**Figure 5.** An illustrative example of the input used for training the cGAN model. The picture on the left shows the source image, which serves as the ground truth or the real data ( $x$ ), as mentioned in the Flow Diagram (Fig. 6 discussed later during the training). The picture on the right represents the simulated II observation pattern of the source (on the left) using a hypothetical observation facility having as array of four IACTs (see Fig. 1) and the tracks of the six telescope baselines (see Fig. 2) generated by these four IACTs due to Earth’s rotation during the observation session. This pattern referred to as  $y$ , in the Flow Diagram (Fig. 6 discussed later) forms the “condition” during the training to which the GAN model has to conform. Salt-and-pepper noise is added to this pattern for enhancing the robustness of the cGAN model. Together, these images form a training pair, where the GAN learns to reconstruct a predicted image (modeling of observed signal) similar to ground truth (left) from the noisy baseline signal (right). The grayscale in both images is normalized to the brightest pixel.

463 At this point, the Discriminator finds its job no better  
 464 than random guessing. A more comprehensive discussion  
 465 of the problem, including proofs, is provided in I.  
 466 Goodfellow et al. (2014).

467 Subsequently, the GAN framework was extended to  
 468 a conditional model (M. Mirza et al. 2014). This new  
 469 model, known as “conditional GAN” or cGAN injects  
 470 a conditioning variable  $y$  into both networks: the Generator  
 471 now generates  $G(z | y)$ , and the Discriminator  
 472 evaluates  $D(x | y)$ . The adversarial objective becomes

$$473 \quad V(D, G) = \mathbb{E}_{x, y \sim p_{\text{data}}(x)} [\log D(x|y)] + \mathbb{E}_{z, y \sim p_z(z)} [\log (1 - D(G(z|y)|y))] \quad (12)$$

474 The conditional variable  $y$  in a cGAN can be various  
 475 additional information including images, labels or text  
 476 contextual to “ground truth” real data  $x$ . Among various  
 477 types of cGANs, Pix2Pix GAN with its image-to-image  
 478 translation design is suitable for the task at hand.  
 479 In our work, we specifically use this conditional variable  
 480 by choosing  $y$  to represent the ground-based intensity-  
 481 interferometry (II) observation patterns: the Generator  
 482 is trained to produce stellar surface images that not  
 483 only look realistic but also conform to the measured II  
 484 correlations, while the Discriminator judges realism *and*  
 485 consistency with the II data.

486 P. Isola et al. (2017) further observed that combining  
 487 the cGAN from Eq. (12) with the traditional  $L_1(G)$  loss  
 488 (also known as the Mean Absolute Error) improves the  
 489 results, as the Generator is encouraged to produce out-  
 490 puts closer to the target image in a pixel-wise sense. We  
 491 adopt this approach in the training of our cGAN model  
 492 by including  $L_1(G)$  defined in eq.(13)

$$493 \quad L_1(G) = \mathbb{E}_{x \sim p_{\text{data}}, z \sim p_z(z)} [\|x - G(z | y)\|_1]. \quad (13)$$

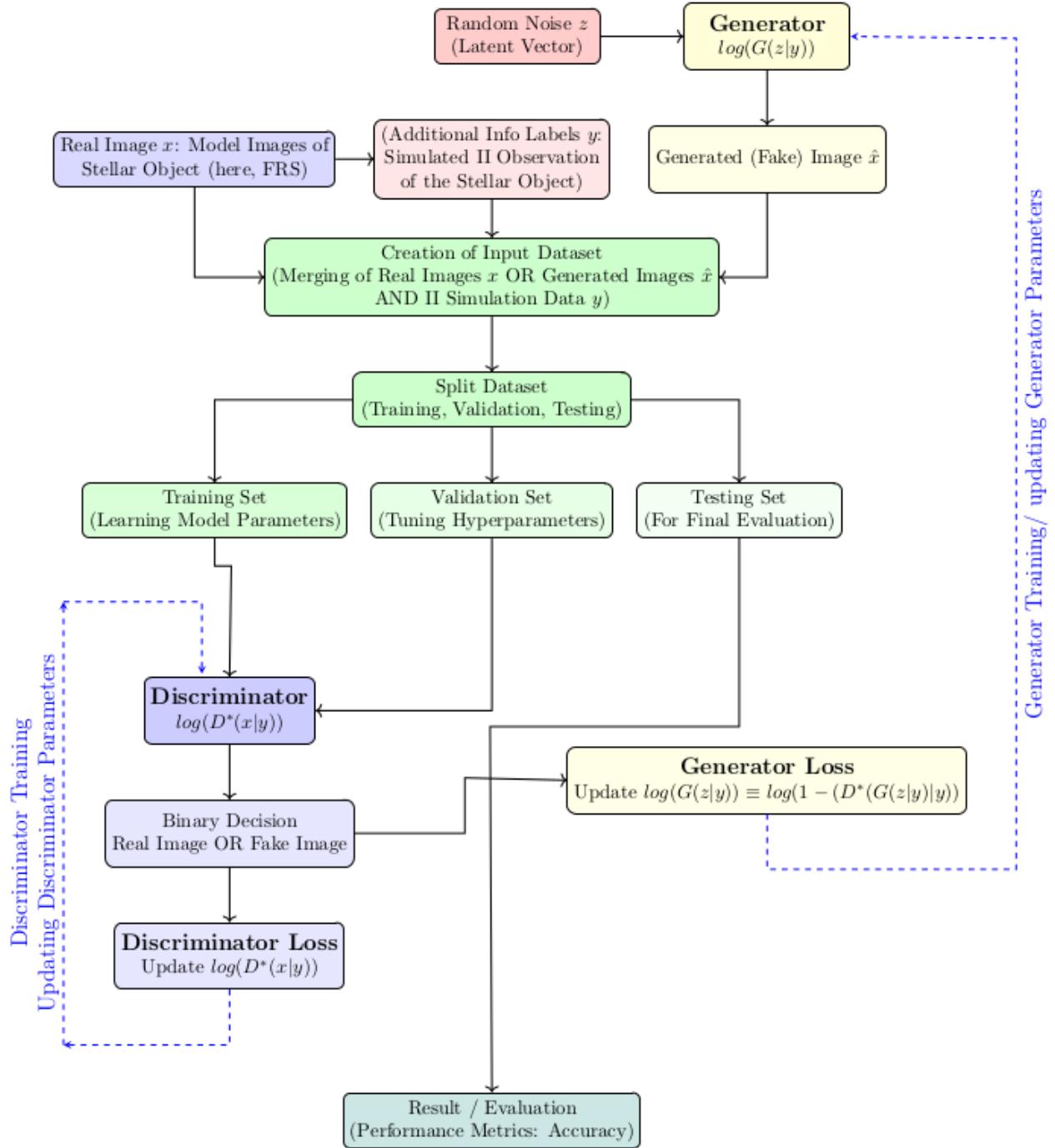
494 The total adversarial loss function, along with the  $L_1$   
 495 loss modulated by a hyperparameter  $\lambda$  then becomes

$$496 \quad L_{\text{tot}} = \arg \min_G \max_D V(D, G) + \lambda \cdot L_1(G). \quad (14)$$

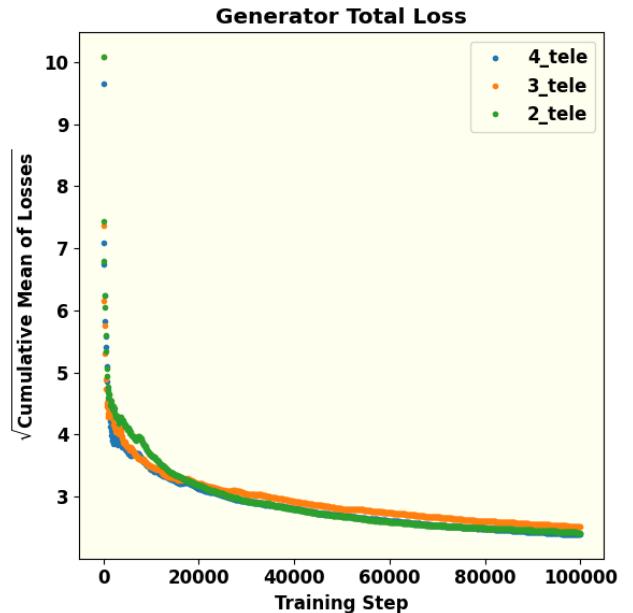
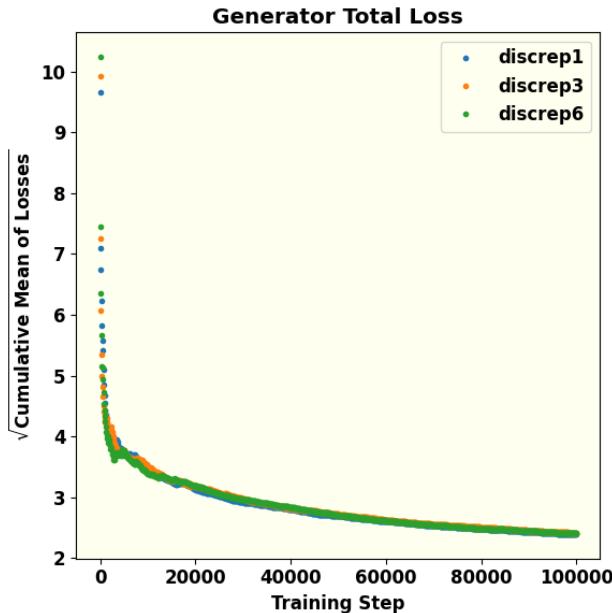
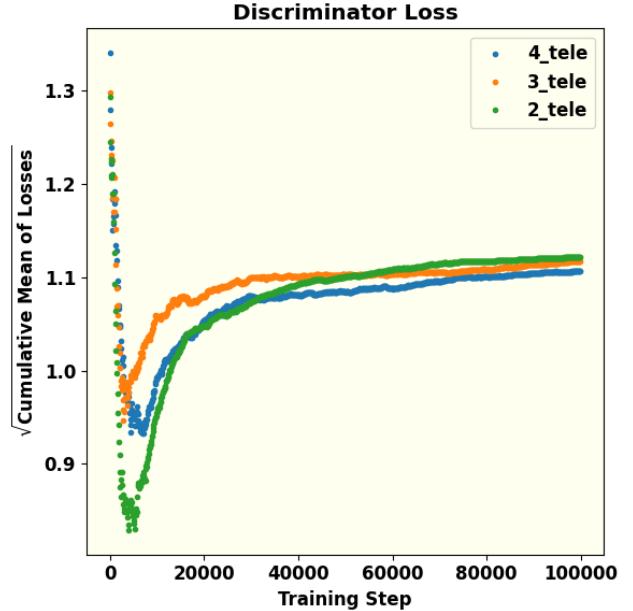
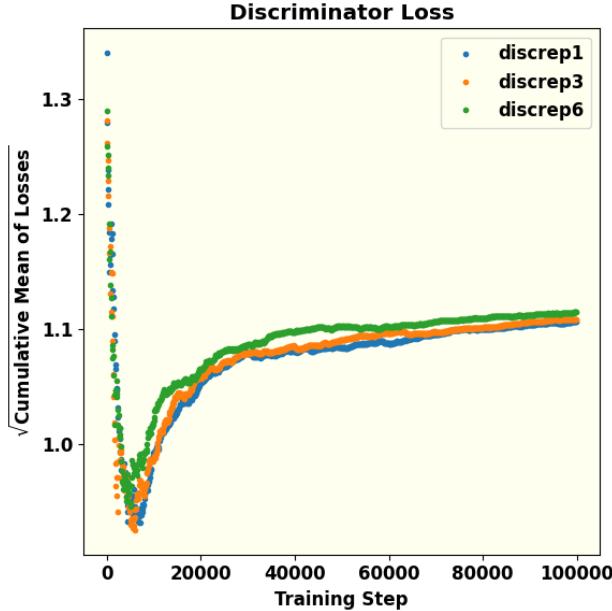
497 This type of network has demonstrated remarkable ro-  
 498 bustness across a variety of applications. For example, it  
 499 can generate colored images from grayscale inputs based  
 500 on architectural labels, transform images from day to  
 501 night, and even predict maps from satellite data. A more  
 502 extensive list of applications is provided in P. Isola et al.  
 503 (2017).

#### 504 4.1. Generator

505 As discussed above, in a GAN the Generator, a deep  
 506 neural network in itself, is responsible for producing  
 507 synthetic data, in this case, images that resemble those



**Figure 6.** A schematic representation of the features of a cGAN model used in this work and the iterative process of its training, validation and testing. The process constitutes four broad stages: (1) choice of an appropriate GAN architecture including both the Discriminator and the Generator (not shown in this figure) (2) preparation of the Training, Validation and Testing datasets and (3) Training and Validation of the Model (4) Testing and Evaluation of the Model. The stages (2), (3) and (4) are depicted in this figure. The datasets are prepared in three broad steps: (i) generating the “ground truth” images of fictitious fast-rotators  $x$ , the sparse II images  $y$  used as the “condition” images in the Model and the generated images  $z$  sampled from a Normal Distribution (ii)merging these images into individual files with  $(x|y)$  and/or  $(z|y)$  (as seen in an illustrated sample in Fig.(5)) and generating the full dataset in this process, and finally (iii) partitioning the full dataset into Training Set, Validation Set and the Testing Set. After the iterative training of the Model and its validation process is complete (“Nash point” of the Minimax Game is reached), the Model is tested using the Testing Set and evaluated.



**Figure 7.** These figures show the effect of the ratio of episodes of Discriminator training per every episode of Generator training. This hyperparameter is termed Discriminator repetition or (discrep) in the figures. The square root of the cumulative mean of the losses are plotted against the training steps for 3 values of this ratio. Understandably, this ratio has a higher impact on the Discriminator loss than the Generator loss. Equal number of episodes of training produces minimum cumulative loss of the Discriminator. The dip in the Discriminator loss during the initial phases of its training can be interpreted as its early success in detecting “fake” (or generated) images because of a poorly trained Generator. With gradual training of the Generator, the success rate of the Discriminator decreases and eventually approaches saturation with equal probability of being successful in telling “fake” from real. The continual decrease and eventual saturation of the Generator loss is a result of its training to generate better images with increasing number of steps.

**Figure 8.** The square root of cumulative mean of Discriminator and Generator loss for different numbers of telescopes. The number of telescopes is also another hyperparameter that has 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 cumulative mean of loss functions is smaller compare to other parameters. For the same reasons as explained in Fig.(7), we observe initial dip and eventual saturation of the Discriminator loss and continual decrease and eventual saturation of the Generator loss with training steps.

of a fast-rotating star. In this work, the Generator is implemented as a U-Net convolutional network (O. Ronneberger et al. 2015). The U-Net consists of a symmetric encoder-decoder structure forming a characteristic “U” shape along with skip connections: the left (contracting or down-sampling) path, the right (expanding or up-sampling path) and the connecting (bottle-neck) path. The left down-sampling path repeatedly applies  $3 \times 3$  convolutions (padded to preserve spatial dimensions) followed by LeakyReLU activations and strided convolution (with stride of 2), by progressively halving the spatial resolution while doubling the channel depth (typically  $64 \rightarrow 128 \rightarrow 256 \rightarrow 512 \rightarrow 1024$ ). At the bottleneck, high-level features are processed without further spatial reduction. The right up-sampling path mirrors this process using  $2 \times 2$  transposed convolutions (stride 2) for up-sampling, halving the channel count at each level and using ReLU as the activation function for all its layers except the output layer. 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). Critically, long skip connections concatenate feature maps from each encoder (down-sampling) level to the corresponding decoder (up-sampling level), directly injecting high-resolution details into the reconstruction process. This enables the network to “remember where everything is” while the deep bottleneck still provides the large receptive field needed to think globally about image structure and semantics. Besides the strided convolutions, modern variants of U-Nets used in state-of-the-art GANs incorporate residual blocks within resolution levels, and frequently add spectral normalization and self-attention at the bottleneck for improved training stability and long-range dependency modelling. These architectural choices allow our U-Net generator to simultaneously recover sharp, high-frequency details (stellar limb edges, limb-darkening profiles, gravity darkening, and rapid-rotation-induced oblateness) and enforce global coherence (overall disk morphology and physical consistency with the observed interferometric visibilities)—making it ideally suited for high-fidelity sky-image reconstruction of fast-rotating stars from sparse ground-based intensity interferometry observations.

During the training, the total Generator Loss function  $L_G$  including the  $L_1$  loss defined in eq.(13) that is minimized is given by

$$\begin{aligned} L_G = & -\mathbb{E}_{z \sim p_z(z)} [\log D(G(z | y) | y)] \\ & + \lambda \mathbb{E}_{x \sim p_{\text{data}}, z \sim p_z(z)} [\|x - G(z | y)\|_1]. \end{aligned} \quad (15)$$

Here,

- $x$  denotes a real data sample (e.g. the “ground-truth” image; here, the synthetically generated fast rotator image) corresponding to condition  $y$ , the simulated II observation data.
- $z$  is a random latent vector drawn from the prior  $p_z(z)$ .
- $G(z | y)$  is the generator output (the reconstructed / synthesized image) given  $z$  and condition  $y$ .
- $D(\cdot | y)$  is the discriminator’s estimate (probability) that its input is “real,” given the same condition  $y$ .
- $\lambda$  is a hyperparameter controlling the trade-off between adversarial realism and pixel-wise fidelity (typical values depend on the problem, e.g. in pix2pix,  $\lambda = 100$ ).

#### 4.2. 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. Each element in the grid corresponds to the “realness” of one patch of the image under examination of the Discriminator at a time. The final loss of the Discriminator is the average over all the patch responses. Evaluating the “realness” of the input image in terms of its constituent patches facilitates capture of texture/ style and other high frequency components in the image. As compared to a global discriminator, it also reduces the number of parameters in the network thereby helping reduce computation cost. It also works on images with arbitrary sizes.

Prior to the down-sampling of data using PatchGAN, each input image is preprocessed with application of Zero Padding followed by batch normalization. The purpose of Zero Padding is to prevent the loss of spatial information and to facilitate the extraction of deeper features from the down-sampled output. Batch normalization is required to stabilize the learning (loss minimization) process. PatchGAN then reduces the spatial dimensions of the images to extract localized features, ensuring the model focuses on smaller regions. In this down-sampling stage, a leaky version of the Rectified Linear Unit (LeakyReLU) is applied in the convolutional

606 layers, similar to the approach used in the Generator.  
 607 The probability  $D(\cdot|y)$  that the patch of the input im-  
 608 age represented by  $\cdot$  is “real” is evaluated through this  
 609 process. The loss function  $D$  of the full input image is  
 610 obtained by averaging over all the patch responses.

611 The Discriminator Loss function  $L_D$  that is optimized  
 612 during the training process is given by

$$613 \quad L_D = -\mathbb{E}_{x \sim p_{\text{data}}(x|y)} [\log D(x|y)] \\ - \mathbb{E}_{z \sim p_z(z)} [\log (1 - D(G(z|y)|y))] \quad (16)$$

614 where the arguments of the  $D$  and  $G$  functions are as  
 615 noted in the text following eq.(15) This loss is composed  
 616 of two parts: one that assesses how accurately it iden-  
 617 tifies fake images (by comparing predictions to a target  
 618 value of 0) and the other that measures how accurately  
 619 the Discriminator identifies real images (by comparing  
 620 predictions to a target value of 1).

621 The training procedure of these two components of  
 622 the cGAN model can be outlined as follows:

- 623 • The discriminator  $D$  is updated by minimizing  
 $L_D$ , keeping  $G$  fixed (so that  $D$  becomes better  
 624 at distinguishing real vs generated images).
- 625 • The generator  $G$  is updated by minimizing  $L_G$ ,  
 626 keeping  $D$  fixed, thus pushing  $G$  to generate im-  
 627 ages that are (a) judged “real” by  $D$ , and (b) close  
 628 (in pixel-wise sense) to the ground-truth  $x$ .

## 630 5. NETWORK PARAMETERS

631 An appropriate cGAN architecture along with a set of  
 632 hyperparameters was optimized by tuning. The objec-  
 633 tive of the model, as already mentioned, was to learn to  
 634 faithfully reproduce a set of sky-images of fast rotators  
 635 subject to the condition that those images are consistent  
 636 with their simulated II observation data. We discuss  
 637 below the architecture and the hyperparameters of the  
 638 cGAN model used for this task. Given the adversarial  
 639 nature of GANs, where the Generator and Discrimina-  
 640 tor engage in a minimax game, careful tuning of key  
 641 parameters is critical to ensure that both networks are  
 642 well-balanced for effective training.

### 643 5.1. Data Preparation

644 First, we generate synthetic images of rapidly rotating  
 645 stars by modelling them as oblate spheroids with vary-  
 646 ing radii and oblateness parameters between 0.5 and 1.  
 647 To incorporate the effect of gravity darkening, we also  
 648 consider different viewing angles and assume a linear de-  
 649 pendence on the declination angle of each point on the  
 650 stellar surface. The traced ellipses result from integrat-  
 651 ing over the source’s hour angle.

652 Next, Salt and Pepper noise is introduced into these  
 653 images; usually this is done at the rate of 0.5% of the  
 654 number of pixels in the image. Then, the images are re-  
 655 sized and their mean is subtracted. A two-dimensional  
 656 Fast Fourier Transform, along with a Fourier shift, is ap-  
 657 plied, yielding a complex number for each pixel. Since II  
 658 does not measure phase, the absolute value is calculated.

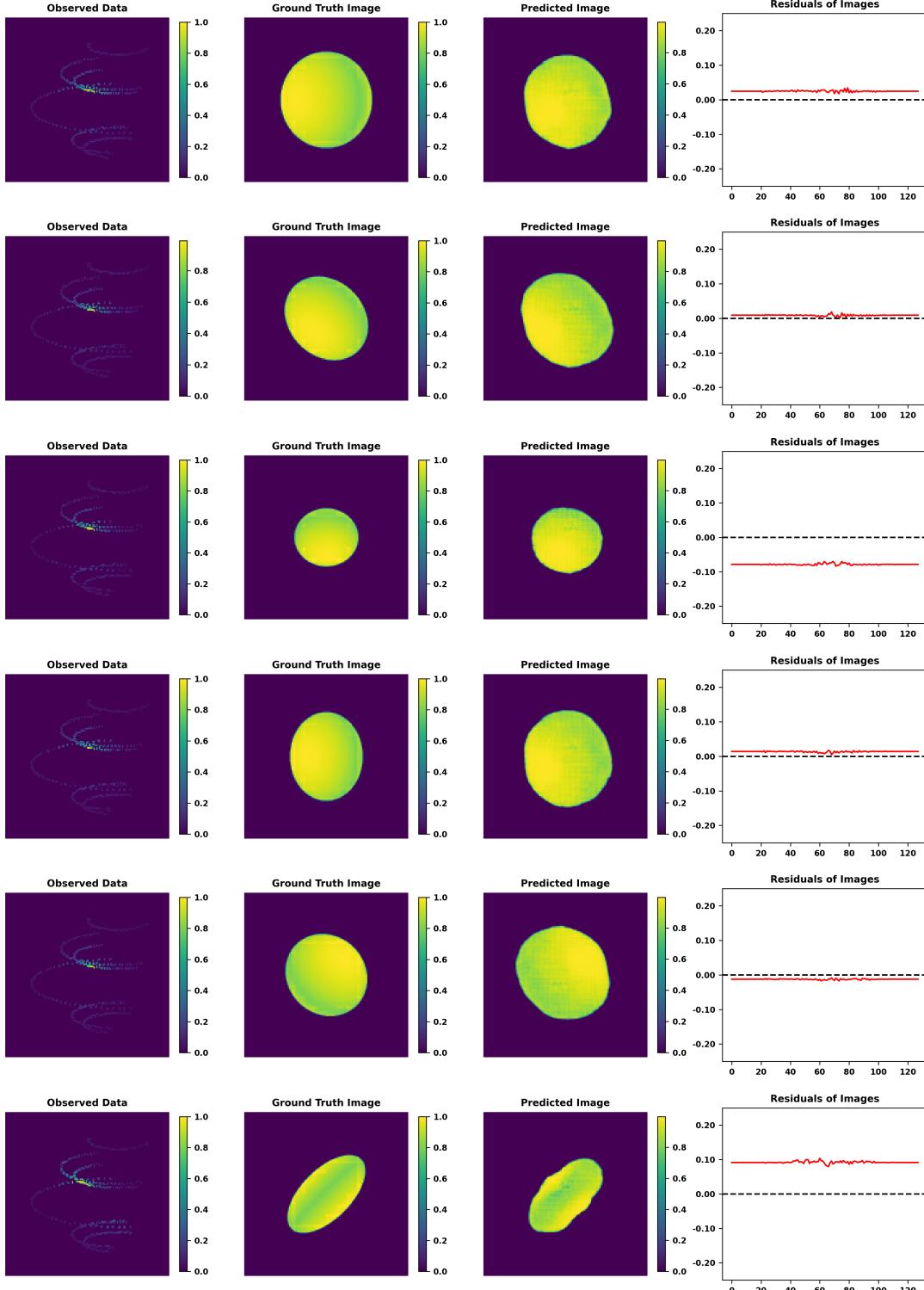
659 Next, sampling of the interferometric plane is in-  
 660 troduced via pixel-wise multiplication of the absolute-  
 661 valued Fourier-transformed image (left panel of Fig. 4)  
 662 and the baseline tracks of (Fig. 2) generated by the four  
 663 IACTs due to Earth rotation during the II observation.  
 664 The result is a map in the Fourier plane featuring several  
 665 ellipses, which is also referred to as the sparse sampling  
 666 map (right panel of fig. 4). This map represents the  
 667 sparse sampling of the signal space corresponding to the  
 668 source (Fig. 3) observed with four telescopes (Fig. 1).  
 669 In essence, the right panel of fig. 4 represents the result  
 670 of simulated II observation of the fictitious fast-rotator  
 671 illustrated in fig. 3.

672 Finally, we normalize the pixel values and convert  
 673 them to 8-bit integers, producing an image that encodes  
 674 the sparsely sampled, phase-free visibility measured by  
 675 II. The image of the corresponding simulated (fictitious)  
 676 star, which serves as the “ground truth” is processed  
 677 identically to avoid any bias. These two images are  
 678 merged side-by-side into a single image (as shown in  
 679 Fig. 5). The dataset thus created is split into three  
 680 parts in the proportions of 80% for training, 10% for  
 681 validation, and 10% as the test set. This partition of  
 682 the full dataset is indicated in the Fig.6. In the follow-  
 683 ing, we refer to these three parts as the Training Set,  
 684 the Validation Set and the Testing Set respectively.

### 685 5.2. GAN Architecture

686 The GAN architecture used in this work is a Pix2Pix  
 687 cGAN, which uses an image-to-image translation strat-  
 688 egy with both the ground truth and the condition being  
 689 images. Originally introduced by Isola et al. P. Isola et  
 690 al. (2017), this architecture is specifically suitable for  
 691 image processing and reconstruction objectives. For in-  
 692 stance, the TensorFlow tutorials<sup>6</sup> demonstrate its appli-  
 693 cation to a dataset of architectural facades. This archi-  
 694 tecture has been adapted for the phase retrieval problem  
 695 at hand here. The network is implemented using the  
 696 TensorFlow library (M. Abadi et al. 2016), calculations  
 697 are performed with scipy (P. Virtanen et al. 2020), and  
 698 plots are generated with matplotlib (J. D. Hunter 2007).

<sup>6</sup> <https://www.tensorflow.org/tutorials/generative/pix2pix>



**Figure 9.** Example results of image reconstruction using the cGAN model along with the II observations simulated in this work. Each row in this figure represents the results for a hypothetical fast-rotating star. The left panel represents the sparse II pattern obtained by the simulated observation of the star using the four IACTs illustrated in Fig.1. This image acts as the “condition” part of the data input to the cGAN model. The second panel from left displays the real image, or ground truth, which the Discriminator uses to distinguish from the images generated by the Generator. The data generated for training, validation and testing of the cGAN model is a merge of this image and its II pattern presented in the left panel. 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.

### 699 5.3. Hyperparameter Tuning

700 The cGAN model architecture used in this work em-  
 701 ploys several hyper-parameters, which are explained  
 702 briefly below (for a more in-depth discussion, see K. P.  
 703 Murphy 2022).

704 The learning rate ( $lr$ ) of the optimizer determines how  
 705 much the model updates its parameters with each iteration. A learning rate that is too small may lead to under-  
 706 fitting, while one that is too large can render the model  
 707 unstable. Therefore, selecting an appropriate learning  
 708 rate is crucial (K. P. Murphy 2022). A canonical choice  
 709 in Pix2Pix and other GAN models is  $lr = 2 \times 10^{-4}$ . In  
 710 our case too, we found this choice to be appropriate.

712 The kernel size refers to the dimensions of the con-  
 713 volutional kernel used in the network, determining how  
 714 many pixels are combined to produce a new pixel. A  
 715 larger kernel size can capture features spanning several  
 716 pixels, but it may also incorporate unrelated features.  
 717 Small kernel sizes are preferable in cases where target  
 718 images finer details or high spatial frequency features.  
 719 Since the “ground truth” target images in our case have  
 720 longer scale gravity darkening features, we have opted  
 721 for the more canonical choice of kernel size being 5.

722 The amount of noise is controlled by two parameters,  
 723 “alpha” and “beta”, which indicate the percentage of  
 724 pixels altered to either white or black, hence the term  
 725 Salt and Pepper noise. Here, “alpha” is applied to the  
 726 real image, while “beta” is applied to the generated im-  
 727 age. Different ratios (“alpha/beta”) can lead to varying  
 728 model performance; however, our results indicate that  
 729 distinct noise rates do not significantly affect the loss  
 730 functions.

731 The batch size defines the number of images processed  
 732 simultaneously by the network. Smaller batch sizes have  
 733 been observed to improve generalization (S. J. Prince  
 734 2023). However, because a larger batch size significantly  
 735 increases training time, a batch size of 1 is used. Besides,  
 736 in Pix2Pix cGAN implementations found in literature  
 737 this choice is found to be often preferred.

738 The buffer size of a Pix2Pix GAN refers to a small  
 739 memory of image pool of previously generated images.  
 740 They are occasionally fed to the Discriminator in place  
 741 of the freshly generated ones. This strategy of mixing  
 742 old and new fakes mitigates the risk of mode collapse  
 743 wherein the Discriminator tends to map all or large  
 744 number of generated images to only one or a few real  
 745 “ground truth” image(s). We have chosen a fairly large  
 746 buffer size (=1400) to protect the model against mode  
 747 collapse.

748 In the training of GANs, one often-followed strategy  
 749 to potentially boost performance is to give the Discrimi-  
 750 nator an advantage by increasing its number of train-

751 ing epochs before returning to the Generator’s training.  
 752 This hyperparameter is referred to as the Discriminator  
 753 repetition (as seen in Table 1). While this can lower the  
 754 Discriminator loss, as shown in Fig. 7, it also increases  
 755 training time. In training our model, we did not notice  
 756 any significant advantage derived from this strategy.  
 757 Since the generated images did not noticeably improve  
 758 with additional Discriminator training, we adopted the  
 759 strategy of training both the networks with equal preference  
 760 (Discriminator repetition = 1).

761 One domain specific hyperparameter of the cGAN  
 762 model presented here is the Number of Telescopes  $N_T$ .  
 763 The degree of sparse sampling of the intensity interfer-  
 764 ometric (II) image plane can be varied to provide the  
 765 model with access to more number of active (non-zero)  
 766 pixels. Point to note here, is that the coverage of the  
 767 Fourier interferometric plane (number of active pixels in  
 768 the II image) scales with available number of baselines,  
 769 and the latter scales quadratically with the number of  
 770 telescopes. Fig. 8 shows the loss functions for different  
 771 numbers of telescopes. There is a significant disparity  
 772 in performance of the model: at the “Nash point”, both  
 773 the loss functions are minimized for the case of four tele-  
 774 scopes. Overall, the degree of sparse sampling appears  
 775 to have the most pronounced effect of all the hyperpa-  
 776 rameters.

777 The hyperparameter Output Channels refers to the  
 778 number of channels in the generator output (e.g., 1 for  
 779 grayscale, 3 for RGB). It is worthwhile to recall that  
 780 the cGAN model constructed in this work is trained on  
 781 “ground truth” target images and the simulated II data,  
 782 both in grey scales as seen in Fig. 5. It is natural that the  
 783 output of this model will be in grey scales only. There-  
 784 fore the value of this hyperparameter is set to 1. The  
 785 choice of hyperparameter  $\lambda$  has been commented upon  
 786 earlier.

787 An optimized set of hyperparameters is selected  
 788 through an iterative process of training and validation.  
 789 For a tentatively chosen set of the hyperparameters,  
 790 the model is trained using the Training Set until the  
 791 both the Discriminator and the Generator loss func-  
 792 tions are minimized. This model, thus trained along  
 793 with its model parameters, is then passed through val-  
 794 idation using the Validation Set. This cycle of training  
 795 and validation is iterated till an optimal set of hyper-  
 796 parameters is arrived at. During each epoch of training  
 797 of the model, both the Discriminator and the Genera-  
 798 tor networks are trained for 100,000 steps. Plots of the  
 799 “Discriminator Loss” function and the “Generator To-  
 800 tal Loss” function presented in Fig. 7 and Fig. 8 represent  
 801 the results of this training for two of the hyperpara-  
 802 meters, namely, the Disriminator repetition (“discrep”, in

short) and the Number of telescopes, respectively. Obviously, the most compute-intensive part of this process is that of the training of the Model. The results of training and validation presented in this work were carried out on a CPU using two nodes, each with 48 threads and the entire process of training and validation required approximately 20 hours on the machine employed for this work. This iterative process of training and validation of the Model is represented schematically in the Fig.6. The chosen optimal set of hyperparameters is presented in the table Table-1. This optimized and trained Model

**Table 1.** Selected hyperparameters used for training the model.

Hyperparameter	Value
Learning rate	2e-4
Kernel size	5
Alpha/Beta	1
Batch size	1
Buffer Size	1400
Discriminator repetition	1
Number of Telescopes	4
Output Channels	1
Lambda	100

is then subjected to testing and evaluation using the Testing Set. The results of this testing and evaluation is presented in the following section.

The Pix2Pix cGAN architecture along with the choice of the values of the hyperparameters mentioned in the Table -1 and used in the implementation of this architecture constitutes the cGAN model (hereinafter referred to as “the GAN Model” or simply “the Model”).

## 6. IMAGE RECONSTRUCTION: RESULTS AND EVALUATION

In this section, we examine the performance of the GAN Model whose architecture and choice of hyperparameters have been discussed above. We subject the trained Model to the task of phase retrieval and image reconstruction on the Testing Dataset.

### 6.1. Visual Evaluation of Images Predicted by the Model

Fig. 9 demonstrates the success of the trained Model in reconstructing the images of a sample of the fictitious fast-rotators drawn from the Testing Set.

The four panels from left to right in each row of Fig. 9 represent the following:

- The left panel represents the sparse II pattern obtained by the simulated observation using the four

IACTs illustrated in Fig.1. As explained earlier, this image acts as part of the input, namely its “condition” part. This image acts as the condition that the sky-images of the star generated by the Generator must conform to.

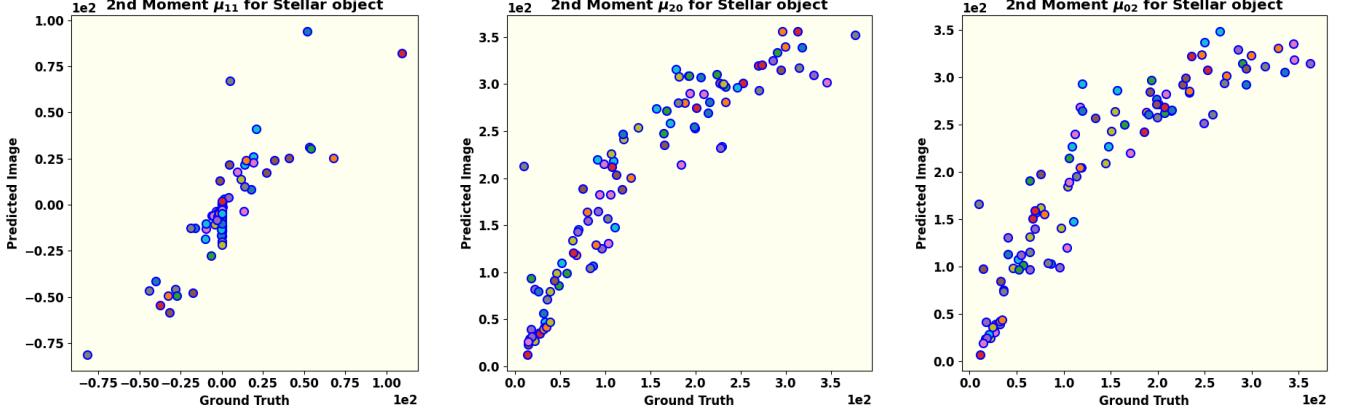
- The second panel from left displays the real image, or ground truth target image, which the Discriminator loss function uses to distinguish from the images generated by the Generator.
- The third panel from left presents the reconstructed (or predicted) image corresponding to the ground truth (second panel) and generated by the trained Model. The similarity of these two images indicates the success of the Model in its stated objective of image reconstruction. We remark that rotating the images by  $180^\circ$  would not change the data. That is, each predicted image contains an arbitrary choice among two possible orientations, differing by  $180^\circ$ .
- The right panel shows the residuals between the ground truth target image and the predicted image in the interferometric plane. The small values are indicative of the success of training the Model.

The predicted images and the residuals presented in the third and the fourth columns (from left) of Fig. 9 show encouraging results, conveying visual information about the source’s size, shape, and brightness distribution across its surface fairly accurately. This has been achieved on the basis of the input provided by II observation using only six baselines, corresponding to four telescopes, at present the maximum on which II is already implemented. Further improvements could surely be achieved by increasing the number of telescopes to enhance the coverage of the  $(u, v)$  plane. A closer examination of this proposition might be able to contribute to the design and instrumentation aspects in the existing and upcoming CTAO.

### 6.2. Evaluation of the Model using Moments

The reconstructed images are visually compelling, demonstrating the Model’s effectiveness in using II to reconstruct images. However, visual assessment alone is insufficient; statistical evaluation of the generated images in comparison with the ground truth images is necessary to validate the results. We present here the results of our calculation and comparison of the moments of distribution pixel brightness in the target “Ground Truth” images and the “Predicted” images.

Image moments capture key properties of the reconstructed objects, such as shape, size, and intensity dis-



**Figure 10.** The second order central moment of the brightness distribution of the images is analogous to the moment of inertia of a mass distribution. The panels of this figure show the comparison of the second order moments ( $\mu_{11}$ ,  $\mu_{20}$ , and  $\mu_{02}$ ) of the predicted images vs. the ground truth images from left to right respectively. Approximate equality of these moments is evident. The small scatter in the moments of the predicted images is indicative of a balanced training of the Model.

tribution, by quantifying features like position, orientation, and brightness distribution. By comparing the moments of the Model-generated images with 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)

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y) \quad (17)$$

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 pixel brightness of the images of the fictitious stars. According to Eq. (17), the monopole of an image is calculated as:

$$M_{00} = \sum_x \sum_y I(x, y). \quad (18)$$

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 first order raw moments normalized by the respective monopole moments of the images given by

$$\begin{aligned} x_c &= \frac{M_{10}}{M_{00}} = \frac{\sum_{x,y} x \cdot I(x, y)}{\sum_{x,y} I(x, y)} \\ y_c &= \frac{M_{01}}{M_{00}} = \frac{\sum_{x,y} y \cdot I(x, y)}{\sum_{x,y} I(x, y)} \end{aligned} \quad (19)$$

represent the centroids  $(x_c, y_c)$  of the pixel brightness distribution of the images.

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:

$$\mu_{pq} = \frac{1}{M_{00}} \sum_x \sum_y (x - x_c)^p (y - y_c)^q I(x, y). \quad (20)$$

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

The second order central moment of the brightness distribution of the images is analogous to the moment of inertia of a mass distribution. Fig. 10 presents the comparison of 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 an approximate equality among these moments, similar to the monopole, thereby confirming the success of applying the Model to reconstruct images with II. The small scatter in the moments of the predicted images indicates robust learning of the Model without over-fitting.

The asymmetry in brightness is calculated using centroid and minimum value of pixels

### 6.3. 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.

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

$$954 \theta = \frac{1}{2} \arctan \left( \frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right). \quad (21)$$

955 The semi-major and semi-minor axes of the stellar ob-  
 956 ject are computed using the second-order central mo-  
 957 ments and are denoted as  $a$  and  $b$ , respectively.

$$958 \begin{aligned} a &= 2\sqrt{mp + \delta} \\ b &= 2\sqrt{mp - \delta} \end{aligned} \quad (22)$$

959 where,

$$960 mp = \frac{\mu_{20} + \mu_{02}}{2} \quad (23)$$

961 and

$$962 \delta = \frac{\sqrt{4\mu_{11}^2 + (\mu_{20} - \mu_{02})^2}}{2}. \quad (24)$$

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

$$965 e = \sqrt{1 - a/b}. \quad (25)$$

966 Eqs. 21-25 describe the elliptical nature of the stellar  
 967 object (in this case, a fast-rotating star) and provide in-  
 968 formation on its shape and size, depending on the com-  
 969 puted values. In contrast, the brightness distribution  
 970 is characterized by skewness, which is quantified using  
 971 third and higher-order moments.

## 972 7. CONCLUSION

973 Intensity Interferometry (II) is re-emerging as a  
 974 promising technique to overcome the challenges of very  
 975 long baseline interferometry in the optical wavelength  
 976 range. However, compared to radio-interferometry, opti-  
 977 cal interferometry faces a major hurdle: photon correla-  
 978 tion captures only the magnitude of the interferometric  
 979 signal, resulting in a loss of phase information.

980 This work addresses the challenge of phase retrieval in  
 981 II using a machine-learning technique, specifically a con-  
 982 ditional Generative Adversarial Network (cGAN). Our  
 983 study demonstrates that the size, shape and brightness  
 984 distribution of fast rotating stars can be recovered by  
 985 a Pix2Pix cGAN model trained on the combined input  
 986 of the sky-image of known sources along with their re-  
 987 spective II data. In this training the sky-image acts  
 988 as the real “ground truth” and the II data acts as the  
 989 “condition”. The Discriminator of our Model is trained  
 990 to efficiently distinguish between the real images and

991 fake (generated) images based on the “ground truth”  
 992 images and the respective II data as the condition. The  
 993 Generator is trained to generate progressively realistic  
 994 images which are also consistent with the condition of  
 995 the II data. The evaluation of the trained Model is then  
 996 carried out by comparison of image moments of ground  
 997 truth images and generated images. Specifically, the  
 998 monopole, second, and third-order moments are com-  
 999 pared. The results support the effectiveness of cGAN in  
 1000 achieving the stated objective.

1001 While the results of this study highlight the significant  
 1002 potential of machine learning, and in particular the ap-  
 1003 plicability of cGAN, for image reconstruction in II, sev-  
 1004 eral aspects require further refinement. First, an impor-  
 1005 tant factor in the reconstruction process is the extent of  
 1006 Fourier plane coverage, which depends on the number of  
 1007 available telescopes and the total observing time. The  
 1008 reasonable success of this piece of work suggests that a  
 1009 network of higher number of telescopes providing higher  
 1010 number of baselines and greater coverage of the  $(u, v)$   
 1011 plane signal, would play a critical role in projects of im-  
 1012 age reconstruction of more complicated stellar systems  
 1013 can be undertaken. Future work might explore different  
 1014 observatory layouts to assess their impact on image re-  
 1015 construction quality. Second, detector efficiencies, which  
 1016 impact the signal-to-noise ratio (SNR) of actual observa-  
 1017 tional data, have not yet been incorporated; addressing  
 1018 these factors will be crucial for more accurate SNR es-  
 1019 timation. Third, exploring and comparing alternative  
 1020 methods for image generation could reveal approaches  
 1021 that outperform cGAN in reconstructing stellar images  
 1022 with II. Fourth, experimenting with different loss func-  
 1023 tions could provide additional insights into the recon-  
 1024 struction quality. Although further testing is needed to  
 1025 refine the GAN and enhance its robustness and reli-  
 1026 ability, our findings suggest that machine learning is a  
 1027 promising approach for phase reconstruction in II.

## 1028 NOTE ON SOFTWARE

1029 The code used for this work is available on the DOI:  
 1030 [10.5281/zenodo.17598807](https://doi.org/10.5281/zenodo.17598807)

## 1031 ACKNOWLEDGEMENTS

1032 One of the authors (SS) gratefully acknowledges the  
 1033 computing facilities and the local hospitality extended  
 1034 to him by the Inter-University Centre for Astronomy  
 1035 and Astrophysics (IUCAA), Pune, India under its Vis-  
 1036 iting Associate Programme during the preparation and  
 1037 finalization of this manuscript.

## REFERENCES

- 1040 Acciari et al., V. A. 2020, MNRAS, 491, 1540  
 1041 Acharyya et al., A. 2024, The Astrophysical Journal, 966,  
 1042 28  
 1043 Aleksić et al., J. 2016, Astroparticle Physics, 72, 76  
 1044 Archer, A., Aufdenberg, J. P., Bangale, P., et al. 2025,  
 1045 arXiv e-prints, arXiv:2506.15027,  
 1046 doi: [10.48550/arXiv.2506.15027](https://doi.org/10.48550/arXiv.2506.15027)  
 1047 Baumgartner et al., S. 2020, MNRAS, 498, 4577  
 1048 Coccioni et al., D. 2021, arXiv preprint arXiv:2122.11578  
 1049 Domiciano de Souza, A., Kervella, P., Jankov, S., et al.  
 1050 2003, Astronomy and Astrophysics, 407, L47,  
 1051 doi: [10.1051/0004-6361:20030868](https://doi.org/10.1051/0004-6361:20030868)  
 1052 Domiciano de Souza, A., Kervella, P., Jankov, S., et al.  
 1053 2005, Astronomy and Astrophysics, 442, 567,  
 1054 doi: [10.1051/0004-6361:20042476](https://doi.org/10.1051/0004-6361:20042476)  
 1055 Dravins et al., D. 2013, Astroparticle Physics, 43, 331,  
 1056 doi: [10.1016/j.astropartphys.2012.04.017](https://doi.org/10.1016/j.astropartphys.2012.04.017)  
 1057 Fienup, J. 1982, Applied Optics, 21, 2758  
 1058 Gamo, H. 1963, Journal of Applied Physics, 34, 875  
 1059 Gerchberg, R. W. 1972, Optik, 35, 237  
 1060 Glauber, R. J. 1963, Physical Review, 130, 2529  
 1061 Goldberger et al., M. L. 1963, Physical Review, 132, 2764  
 1062 Goodfellow et al., I. 2014, Advances in neural information  
 1063 processing systems, 27  
 1064 Hanbury Brown, R., & Twiss, R. Q. 1956, Nature, 177, 27  
 1065 Hanbury Brown et al., R. 1957, Proceedings of the Royal  
 1066 Society of London. Series A. Mathematical and Physical  
 1067 Sciences, 242, 300  
 1068 Hanbury Brown et al., R. 1958, Proceedings of the Royal  
 1069 Society of London. Series A. Mathematical and Physical  
 1070 Sciences, 243, 291  
 1071 Hanbury Brown et al., R. 1974, Monthly Notices of the  
 1072 Royal Astronomical Society, 167, 121  
 1073 Haubois et al., X. 2009, Astronomy & Astrophysics, 508,  
 1074 923, doi: [10.1051/0004-6361/200912927](https://doi.org/10.1051/0004-6361/200912927)  
 1075 Holmes et al., R. 2010, in Adaptive Coded Aperture  
 1076 Imaging, Non-Imaging, and Unconventional Imaging  
 1077 Sensor Systems II, Vol. 7818, SPIE, 175–185  
 1078 Hu, M.-K. 1962, IRE transactions on information theory, 8,  
 1079 179  
 1080 Hunter, J. D. 2007, Computing in Science & Engineering, 9,  
 1081 90, doi: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55)  
 1082 Isola et al., P. 2017, in Proceedings of the IEEE conference  
 1083 on computer vision and pattern recognition, 1125–1134  
 1084 Kirisits et al., C. 2024, arXiv preprint arXiv:2406.14143v2  
 1085 Le Bohec, S.; Holder, J. 2006, The Astrophysical Journal,  
 1086 649, 399  
 1087 Li et al., X. 2014, in Imaging Spectroscopy, Telescopes and  
 1088 Large Optics, Vol. 9298, SPIE, 92981G1–G7  
 1089 Liu, L.-C., Wu, C., Li, W., et al. 2024, arXiv preprint,  
 1090 arXiv:2404.15685, doi: [10.48550/arXiv.2404.15685](https://doi.org/10.48550/arXiv.2404.15685)  
 1091 Liu et al., L.-C. 2025, Physical Review Letters, 134, 180201,  
 1092 doi: [10.1103/PhysRevLett.134.180201](https://doi.org/10.1103/PhysRevLett.134.180201)  
 1093 Lucy, L. B. 1967, Zeitschrift für Astrophysik, Vol. 65, p. 89,  
 1094 65, 89  
 1095 Maeder, A. 1999, Astronomy & Astrophysics, 347, 185  
 1096 Martinez, A. O., Baron, F., Monnier, J. D., Roettenbacher,  
 1097 R. M., & Parks, J. R. 2021, The Astrophysical Journal,  
 1098 916, 60, doi: [10.3847/1538-4357/ac06a5](https://doi.org/10.3847/1538-4357/ac06a5)  
 1099 McAlister et al., H. A. 2005, The Astrophysical Journal,  
 1100 628, 439  
 1101 Mirza et al., M. 2014, arXiv preprint arXiv:1411.1784  
 1102 Monnier, J. D., Zhao, M., Pedretti, E., et al. 2007, Science,  
 1103 317, 342, doi: [10.1126/science.1143205](https://doi.org/10.1126/science.1143205)  
 1104 Murphy, K. P. 2022, Probabilistic machine learning: an  
 1105 introduction (MIT press)  
 1106 Mustafa et al., M. 2019, Computational Astrophysics and  
 1107 Cosmology, 6, 1  
 1108 Norris et al., R. P. 2021, The Astrophysical Journal, 919,  
 1109 124, doi: [10.3847/1538-4357/ac0c7e](https://doi.org/10.3847/1538-4357/ac0c7e)  
 1110 Nuñez et al., P. D. 2010, in Optical and Infrared  
 1111 Interferometry II, Vol. 7734, SPIE, 458–467  
 1112 Nuñez et al., P. D. 2012a, Monthly Notices of the Royal  
 1113 Astronomical Society, 419, 172  
 1114 Nuñez et al., P. D. 2012b, Monthly Notices of the Royal  
 1115 Astronomical Society, 424, 1006  
 1116 Pedretti, E., Monnier, J. D., ten Brummelaar, T., &  
 1117 Thureau, N. D. 2009, New Astronomy Reviews, 53, 353,  
 1118 doi: [10.1016/j.newar.2010.07.005](https://doi.org/10.1016/j.newar.2010.07.005)  
 1119 Prince, S. J. 2023, Understanding deep learning (MIT  
 1120 press)  
 1121 Rai, K. N., Basak, S., Sarangi, S., & Saha, P. 2025,  
 1122 Resonance, 30, 45, doi: [10.1007/s12045-025-1729-x](https://doi.org/10.1007/s12045-025-1729-x)  
 1123 Rai et al., K. N. 2021, Monthly Notices of the Royal  
 1124 Astronomical Society, 507, 2813,  
 1125 doi: [10.1093/mnras/stab2391](https://doi.org/10.1093/mnras/stab2391)  
 1126 Rai et al., K. N. 2022, Monthly Notices of the Royal  
 1127 Astronomical Society, 516, 2864,  
 1128 doi: [10.1093/mnras/stac2433](https://doi.org/10.1093/mnras/stac2433)  
 1129 Ronneberger et al., O. 2015, in Medical image computing  
 1130 and computer-assisted intervention—MICCAI 2015: 18th  
 1131 international conference, Munich, Germany, October 5–9,  
 1132 2015, proceedings, part III 18, Springer, 234–241  
 1133 Sato et al., T. 1978, Applied optics, 17, 2047  
 1134 Sato et al., T. 1979, Applied Optics, 18, 485  
 1135 Sato et al., T. 1981, Applied Optics, 20, 2055  
 1136 Schawinski et al., K. 2017, MNRAS, 467, L110  
 1137 Teague, M. R. 1980, Josa, 70, 920

- <sup>1138</sup> Teague, M. R. 1983, Journal of Opticala Society of  
<sup>1139</sup> America, 73, 1434
- <sup>1140</sup> van Belle, G. T., Ciardi, D. R., Thompson, R. R., Akeson,  
<sup>1141</sup> R. L., & Lada, E. A. 2001, The Astrophysical Journal,  
<sup>1142</sup> 559, 1155, doi: [10.1086/322340](https://doi.org/10.1086/322340)
- <sup>1143</sup> Virtanen et al., P. 2020, Nature methods, 17, 261
- <sup>1144</sup> Vogel et al., N. 2025, MNRAS, 537, 2334
- <sup>1145</sup> Von Zeipel, H. 1924, Monthly Notices of the Royal  
Astronomical Society, Vol. 84, p. 665-683, 84, 665
- <sup>1146</sup> Zhang et al., J. 2020, Opt. Lett., 45, 3649
- <sup>1147</sup> Zhao, M., Monnier, J. D., Pedretti, E., et al. 2009, The  
Astrophysical Journal, 701, 209,  
<sup>1149</sup> doi: [10.1088/0004-637X/701/1/209](https://doi.org/10.1088/0004-637X/701/1/209)
- <sup>1150</sup>