

## 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.

### 20 1. INTRODUCTION

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

43 The foundational basis of II stems from the pioneering  
44 experiments and theoretical investigations, widely  
45 referred to as the “HBT effect”, carried out initially by

46 Hanbury Brown and Twiss (HBT) R. Hanbury Brown  
47 & R. Q. Twiss (1956); R. Hanbury Brown et al. (1957,  
48 1958) and, later, by Glauber (R. J. Glauber 1963). The  
49 correlation between photons (called HBT effect) measured  
50 by a pair of photon detectors in two coherent  
51 beams of light was reported in 1956 R. Hanbury Brown  
52 & R. Q. Twiss (1956). K. N. Rai et al. (2025) present  
53 a modern-day variant of this experiment, carried out  
54 with pseudo-thermal light. The HBT Effect and the related  
55 theoretical investigations laid the foundation for  
56 the modern field of Quantum Optics.

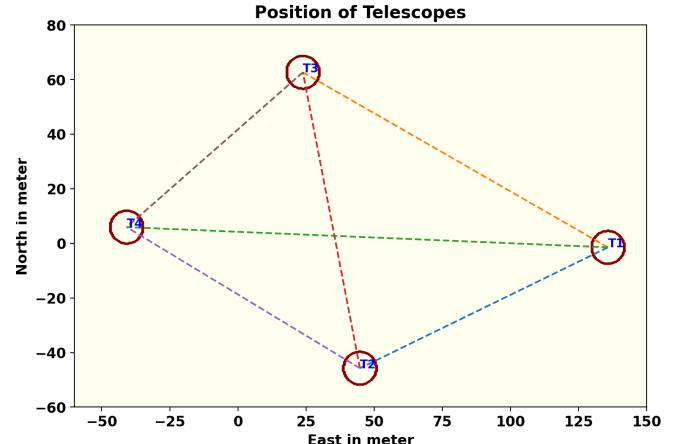
57 Hanbury Brown and his collaborators led the creation  
58 and installation of the iconic II facility at Narrabri, Australia  
59 and reported the measurement of angular diameters of 32 stars and a few multiple star-systems (R. Hanbury Brown et al. 1974). Soon after this work, however, II observations 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 HESS are now being reported (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 carried out by Rai et al. (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. The 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 in a nascent stage. The 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, and especially, that by A. Archer et al. (2025) put our work in con-



**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 Intensity Interferometry (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.

text. Thus our work presented here is a natural next step of the work of Archer et al and others. We emulate the 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.

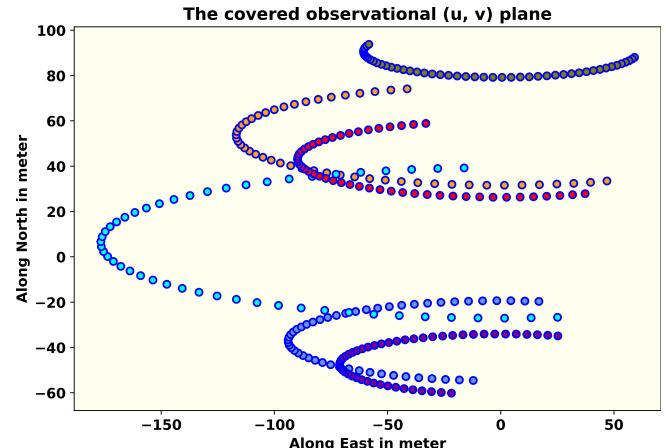
This paper is organized as follows. The next section discusses briefly the past efforts at image reconstruction on II, followed by the section on a discussion of Intensity Interferometry, focusing on its signal and noise characteristics for fast-rotating stars along the Earth's rotation. The following 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.

151            2. PAST EFFORTS AT IMAGE  
 152            RECONSTRUCTION IN INTENSITY  
 153            INTERFEROMETRY

154     Several approaches have been developed for phase re-  
 155     construction in intensity interferometry (II). H. Gamo  
 156     (1963) introduced triple-intensity correlation, applied by  
 157     M. L. Goldberger et al. (1963) to microscopic systems  
 158     and extended by T. Sato et al. (1978, 1979, 1981) to  
 159     measure stellar diameters and phases, though limited  
 160     by low signal-to-noise ratio (SNR). R. W. Gerchberg  
 161     (1972) proposed an iterative phase retrieval method us-  
 162     ing image and diffraction plane data, sensitive to initial  
 163     estimates and convergence speed. J. Fienup (1982) im-  
 164     proved this with the Hybrid Input-Output algorithm,  
 165     enhancing robustness in noisy conditions. R. Holmes et  
 166     al. (2010) utilized Cauchy-Riemann relations for 1-D  
 167     and 2-D image reconstruction, applied to simulated  
 168     stellar data with Imaging Cherenkov Telescope Arrays  
 169     (P. D. Nuñez et al. 2010, 2012a,b), but faced compu-  
 170     tational complexity for higher dimensions. X. Li et al.  
 171     (2014) introduced a regularized iterative method incor-  
 172     porating priors (e.g., sparsity) to mitigate noise and  
 173     ill-posedness, though challenged by parameter tuning  
 174     and initial guess sensitivity. The Transport-of-Intensity  
 175     Equation (TIE), proposed by M. R. Teague (1983), re-  
 176     trievals phase from intensity variations across planes;  
 177     J. Zhang et al. (2020) solved TIE as a Poisson equa-  
 178     tion using a maximum intensity assumption, while C.  
 179     Kirisits et al. (2024) combined TIE with the Trans-  
 180     port of Phase Equation for improved accuracy across  
 181     arbitrary apertures and non-uniform illumination, with  
 182     convergence dependent on initial guesses and boundary  
 183     conditions.

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

200     Furthermore, given that achieving a high signal-to-  
 201     noise ratio (SNR) is often challenging in astronomical  
 202     datasets, it is immensely beneficial if additional data can



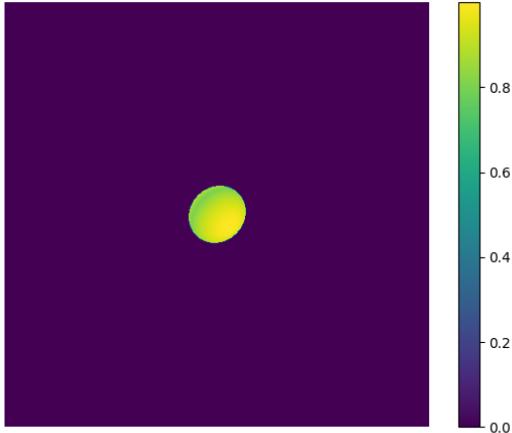
203     **Figure 2.** The tracks of the baselines due to Earth ro-  
 204     tation (described in sec. 3.3) provided by the hypothetical  
 205     observation facility of four telescopes arranged in fig. 1 for  
 206     one night of observation. The number of baselines scales as  
 207     square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

210     be generated using the available information from the  
 211     observed sky density distribution and ground-based ob-  
 212     servations (II data, in our case) of the sources under in-  
 213     vestigation. Generative Adversarial Networks (GANs),  
 214     introduced by I. Goodfellow et al. (2014), have been  
 215     successful in such data augmentation tasks. Conditional  
 216     GAN (cGAN) architectures, proposed by M. Mirza et al.  
 217     (2014) and applied to a wide variety of datasets by P.  
 218     Isola et al. (2017), leverage additional information about  
 219     the images in the training datasets and have demon-  
 220     strated remarkable robustness in image recovery across  
 221     diverse data types.

222     In the astrophysical context, K. Schawinski et al.  
 223     (2017) employed a GAN model to recover features —  
 224     such as spiral arms, central bulges, and disk structures  
 225     of galaxies — from noise-affected images. M. Mustafa et  
 226     al. (2019) developed and customized a Deep Convo-  
 227     lutional GAN, dubbed “CosmoGAN”, capable of gen-  
 228     erating high-fidelity weak-lensing convergence maps of  
 229     dark matter distribution that statistically reproduce real  
 230     weak lensing structures. D. Coccolini et al. (2021)  
 231     have successfully generated credible images of planets,  
 232     nebulae, and galaxies using “lightweight” and “physics-  
 233     uninformed” GANs to produce synthetic images of cele-  
 234     stial bodies. They also generated a “Hubble Deep Field-  
 235     inspired” wide-view simulation of the universe.

236            3. INTENSITY INTERFEROMETRY (II) WITH  
 237            IACT ARRAYS

238     This section presents a brief conceptual overview of  
 239     how an array of telescopes is used to perform II obser-



**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.

vations, and explains the Signal-to-Noise Ratio (SNR) from these measurements.

### 3.1. 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-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.

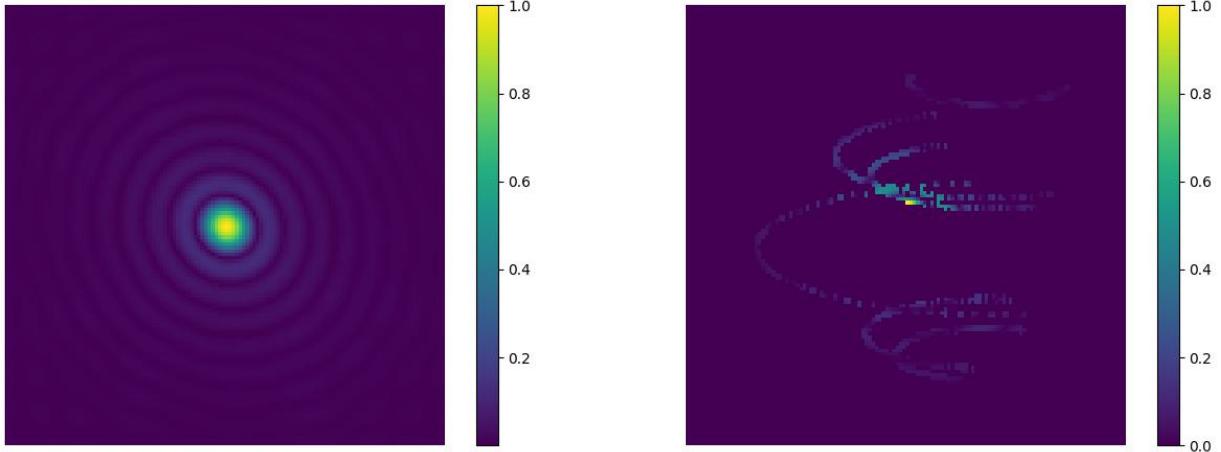
### 3.2. The Signal-to-Noise Ratio for II

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).



**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.

### 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  $b$ .

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

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. Dravins 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 van Belle et al 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 Archer et al. 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.

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. (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. 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

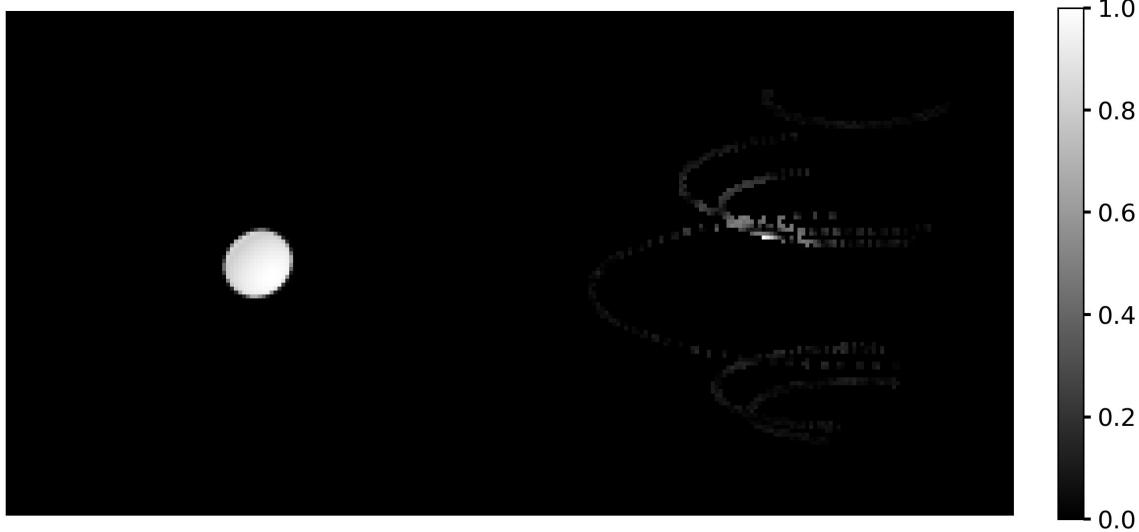
with baselines as seen in Fig. 1) over one night. Using this modest amount of signal from one night's observation, we have trained a cGAN to construct the image of the source.

#### 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 the 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  $\theta_G$



**Figure 5.** 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 Intensity Interferometry (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.

and  $\theta_D$  are optimized such that  $p_G(x_{gen})$  gets maximally aligned with  $p_{data}(x)$ .

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

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

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  $p_G = p_{data}$ . 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  $p_{data}$ .

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

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

Subsequently, the GAN framework was extended to a conditional model (M. Mirza et al. 2014). This new model, known as “conditional GAN” or cGAN injects a conditioning variable  $y$  into both networks: the Generator now generates  $G(z | y)$ , and the Discriminator

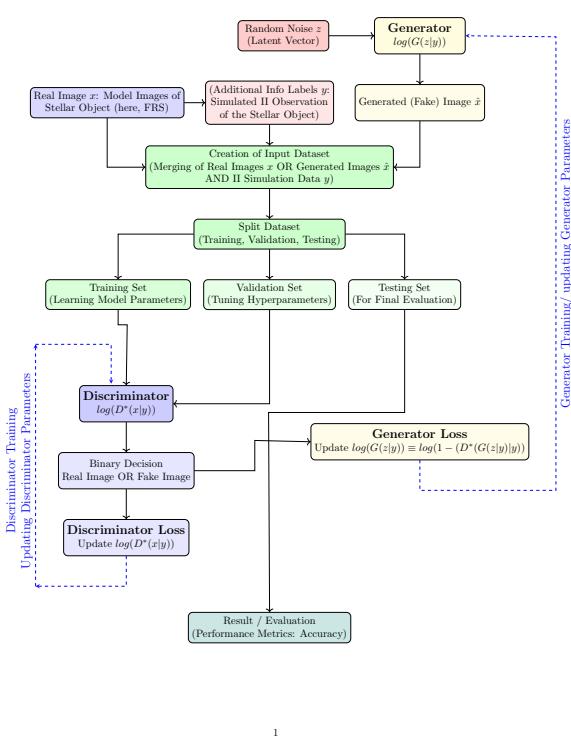
evaluates  $D(x | y)$ . The adversarial objective becomes

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

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

P. Isola et al. (2017) further observed that combining the cGAN from Eq. (12) with the traditional  $L_1(G)$  loss (also known as the Mean Absolute Error) improves the results, as the Generator is encouraged to produce outputs closer to the target image in a pixel-wise sense. We adopt this approach in the training of our cGAN model by including  $L_1(G)$  along with the hyperparameter  $\lambda = 100$ . The total loss function that is minimized is:

$$L_{tot} = \arg \min_G \max_D V(D, G) + \lambda \cdot L_1(G) \quad (13)$$

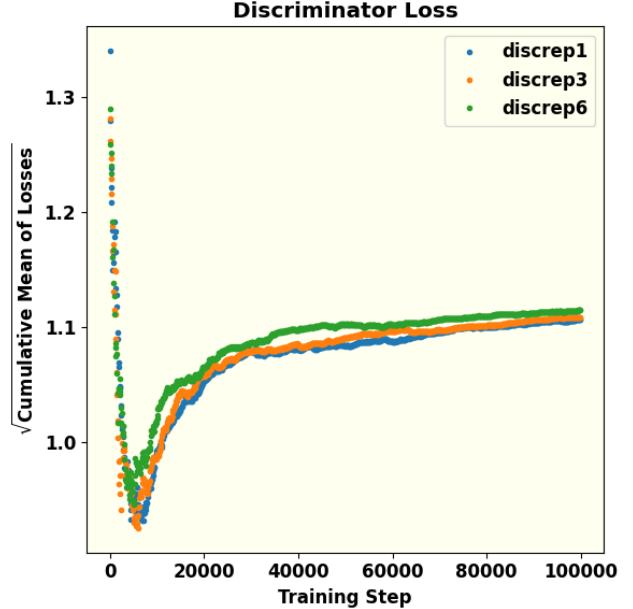


**Figure 6.** A schematic representation of the cGAN model used in this work. The diagram depicts the features of cGAN: (1) the inputs of the real “ground truth” images  $x$ , the conditional images  $y$  and the generated image  $z$  sampled from a Normal Distribution (2) the creation of the training data of the Discriminator network (3) training of the Discriminator and the Generator networks and optimization of their respective Loss functions. The data sets used in training, validation and testing of the networks are sets of merged images  $(x|y)$  and  $(z|y)$  where  $x$  and  $y$  represent the source image of the simulated fictitious star (illustrated in Fig.3) and its sparse ground observation Intensity Interferometry (II) image (illustrated in the right panel of Fig.(4)). These data sets are generated by mixing the merged  $(x|y)$  and  $(z|y)$  images in equal proportions. After the training of both the networks is complete (“Nash point” of the Minimax Game is reached), the success of the model is tested using the testing data.

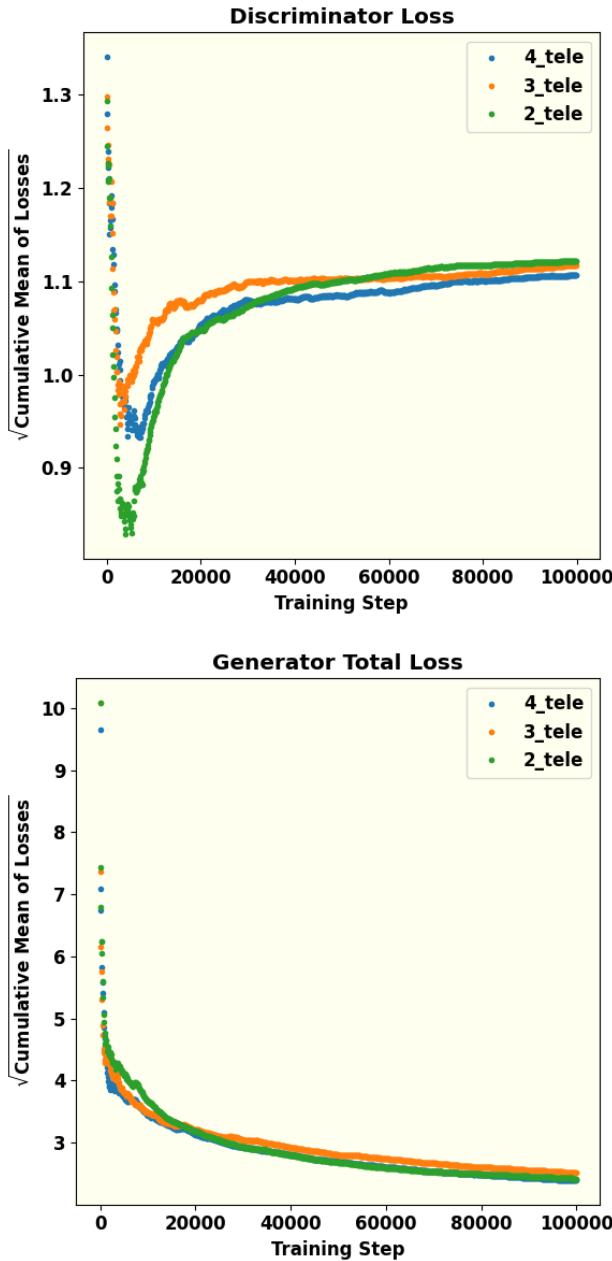
502 with the choice  $\lambda = 100$  and

$$L_1(G) = \mathbb{E}_{x,y,z} [||y - G(x, z)||_1] \quad (14)$$

503 This type of network has demonstrated remarkable robustness across a variety of applications. For example, it  
 504 can generate colored images from grayscale inputs based  
 505 on architectural labels, transform images from day to  
 506 night, and even predict maps from satellite data. A more  
 507 extensive list of applications is provided in P. Isola et al.  
 508 (2017).



**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.

#### 4.1. Generator

As discussed above, in a GAN the Generator, a deep neural network in itself, 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). 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.

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.

#### 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 is 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.**

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 is fed to the Generator 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.

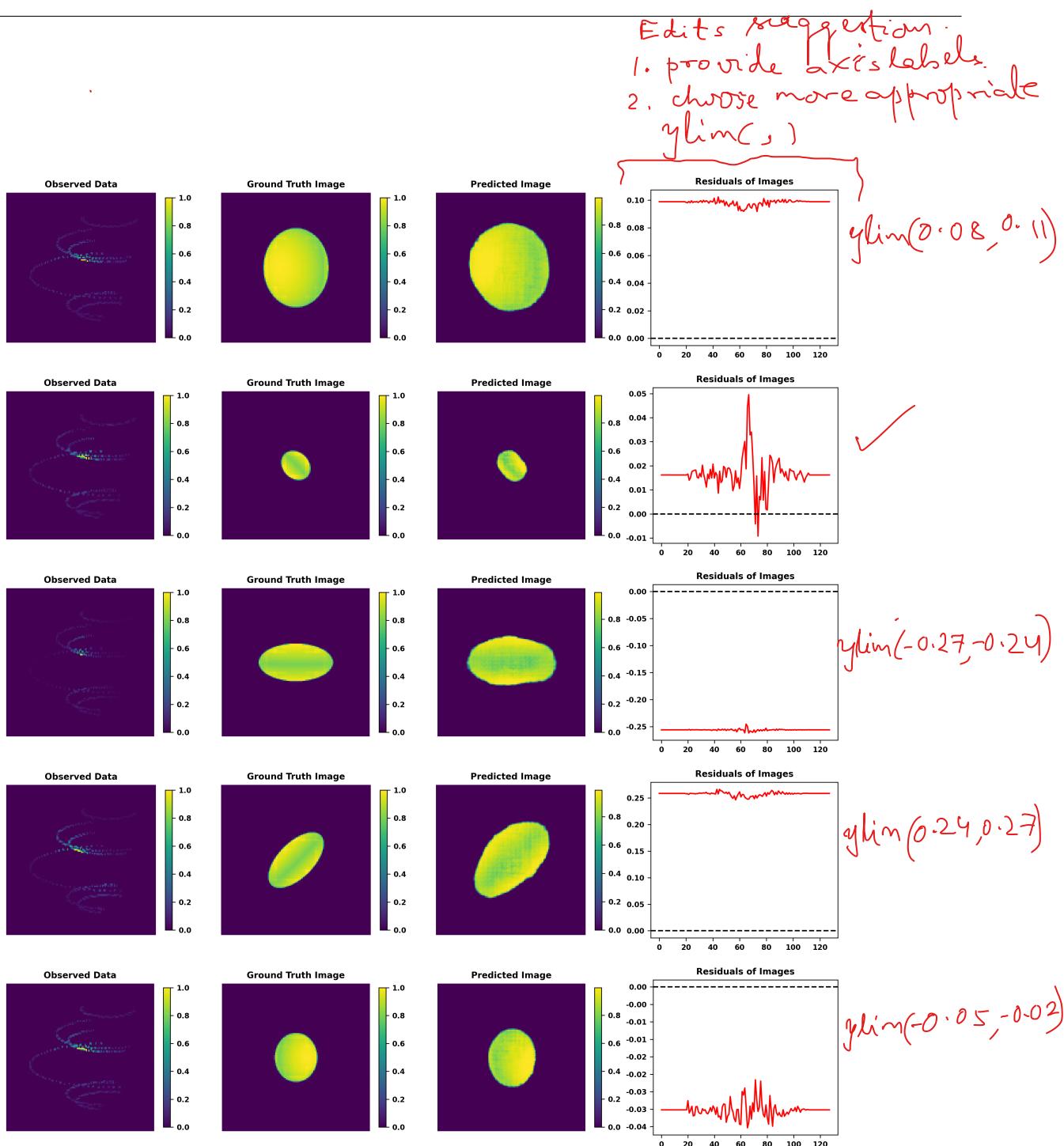
### 5. NETWORK PARAMETERS

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

#### 5.1. Data Preparation

First, we simulate generate fictitious images of 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 on declination angle on the stellar surface for generating 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 into these images; usually this is done at the rate of 0.5% of the number of pixels in the image. 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.



**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. The uniform background color in the difference panel suggests that the network is picking up white noise from the  $(u, v)$  plane.

Discussion and rewriting needed.

665     **Question to discuss: Why are we calculating  
666     FFT of the images??? These images are by them-  
667     selves the Fourier Transform of the sky image of  
668     the source. What do we obtain by calculating  
669     their FFT???**

670     Next, sparse sampling of the interferometric plane is  
671     introduced via pixel-wise multiplication of the absolute-  
672     valued Fourier-transformed image (left panel of Fig. 4)  
673     and the sparse sampling map baseline tracks of (Fig. 2)  
674     generated by the four IACTs due to Earth rotation during  
675     the II observation. The result is a map in the Fourier  
676     plane featuring several ellipses, which is also referred to  
677     as the sparse sampling map (right panel of fig. 4). This  
678     map represents the sparse sampling of the signal space  
679     corresponding to the source (Fig. 3) observed with four  
680     telescopes (Fig. 1). In essence, the right panel of fig. 4  
681     represents the result of simulated II observation of the  
682     fictitious fast-rotator illustrated in fig. 3.

683     Finally, we normalize the pixel values and convert  
684     them to 8-bit integers, producing an image that encodes  
685     the sparsely sampled, phase-free visibility measured by  
686     II. The image of the corresponding simulated (fictitious)  
687     star, which serves as the “ground truth” is processed  
688     identically to avoid any bias. These two images are  
689     merged side-by-side into a single image (as shown in  
690     Fig. 5). The dataset thus created is split into three  
691     parts in the proportions of 80% for training, 10% for  
692     validation, and 10% as a held-out test set.

### 693     5.2. GAN Architecture

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

### 708     5.3. Hyperparameter Tuning

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

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

713     The learning rate of the optimizer determines how  
714     much the model updates its parameters with each it-  
715     eration. A learning rate that is too small may lead  
716     to underfitting, while one that is too large can render  
717     the model unstable. Therefore, selecting an appropriate  
718     learning rate is crucial (K. P. Murphy 2022).

719     The kernel size refers to the dimensions of the con-  
720     volutional kernel used in the network, determining how  
721     many pixels are combined to produce a new pixel. A  
722     larger kernel size can capture features spanning several  
723     pixels, but it may also incorporate unrelated features.

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

733     The batch size defines the number of images processed  
734     simultaneously by the network. Smaller batch sizes have  
735     been observed to improve generalization (S. J. Prince  
736     2023). However, because a larger batch size significantly  
737     increases training time, a batch size of 1 is used.

738     When training GANs, one strategy to potentially  
739     boost performance is to give the Discriminator an ad-  
740     vantage by increasing its number of training steps before  
741     returning to the Generator’s training. While this can  
742     lower the Discriminator loss, as shown in Fig. 7, it also  
743     increases training time and leads to a slight rise in the  
744     Generator loss. In training our model, we did not no-  
745     tice any significant advantage derived from this strategy.  
746     Since the generated images do not noticeably improve  
747     with additional Discriminator training, both networks  
748     are typically trained with the same number of steps.

749     Finally, the degree of sparse sampling can be varied  
750     to provide the model with access to more active (non-  
751     zero) pixels. Increasing the number of telescopes results  
752     in more baselines and, consequently, more available pix-  
753     els. Fig. 8 shows the loss functions for different numbers  
754     of telescopes. There is a significant disparity in perfor-  
755     mance: at “Nash point”, both the loss functions are  
756     minimized for the case of four telescopes. This is partly  
757     because the relationship between telescopes and base-  
758     lines is quadratic: the coverage of the Fourier plane of  
759     Intensity Interferometry (II) scales as square of the num-  
760     ber of telescopes employed in the II observation. For  
761     example, the Fourier plane can be sampled along six  
762     tracks when using four telescopes, as compared to only  
763     one track if using only two telescopes. In the case of  
764     two telescopes, both the Generator and Discriminator

765 exhibit less smooth training, as indicated by the outliers. Performance improves with three telescopes and  
 766 becomes very promising with four. Overall, the degree  
 767 of sparse sampling appears to have the most pronounced  
 768 effect of all the hyperparameters.  
 769

770 All the selected hyperparameters to train the model  
 771 is summarized in the given table:

**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
No. of Telescope	4
Output Channels	1
Lambda	100

- 802 The second panel from left displays the real image,  
 803 or ground truth, which the Discriminator loss func-  
 804 tion uses to distinguish from the images generated  
 805 by the Generator. During training, the GAN aims  
 806 to keep minimum difference between these ground  
 807 truth image and generated images. The data gen-  
 808 erated for training, validation and testing of the  
 809 cGAN model is a merge of this image and its II  
 810 pattern presented in the left panel.

- 811 The third panel from left presents the recon-  
 812 structed, or predicted, image produced by the  
 813 trained GAN, highlighting its success in image re-  
 814 construction.
- 815 • **Check if correct:** The right panel shows the  
 816 difference between the ground truth and the  
 817 predicted image in the interferometric  
 818 plane. The uniform background color of the  
 819 frames suggests that the GAN model picks  
 820 up white noise in the interferometric plane.  
 821 I also think that axis labels need to be pro-  
 822 vided.

823 The predicted images presented in the third column  
 824 (from left) of Fig. 9 show encouraging results, ac-  
 825 curately conveying visual information about the source's  
 826 size, shape, and brightness distribution across its surface  
 827 fairly accurately. This has been achieved on the basis of  
 828 the input provided by II observation using only six base-  
 829 lines. However, Arguably, further improvements can be  
 830 achieved by increasing the number of telescopes to max-  
 831 imize coverage of the  $(u, v)$  plane. A closer examination  
 832 of this proposition might be able to contribute to the  
 833 design and instrumentation aspects in the existing and  
 834 upcoming CTAO. an ideal candidate for this approach.

## 835 6.2. Evaluation of GAN using Moments

836 The reconstructed images are visually compelling,  
 837 demonstrating the GAN's effectiveness in using II to re-  
 838 construct images. However, visual assessment alone is  
 839 insufficient; statistical evaluation is necessary to validate  
 840 the results. To achieve this, we employ image moments  
 841 as a statistical method. Image moments capture key  
 842 properties of the reconstructed objects—such as shape,  
 843 size, and intensity distribution—by quantifying features  
 844 like position, orientation, and brightness distribution.  
 845 By comparing the moments of the GAN-generated im-  
 846 ages to those of the ground truth, we can objectively as-  
 847 sess the consistency and accuracy of the reconstruction.  
 848 This approach provides a reliable framework for evaluat-  
 849 ing reconstruction quality, as image moments can reveal  
 850 subtle differences in geometric and intensity properties

## 772 6. IMAGE RECONSTRUCTION: RESULTS AND 773 EVALUATION

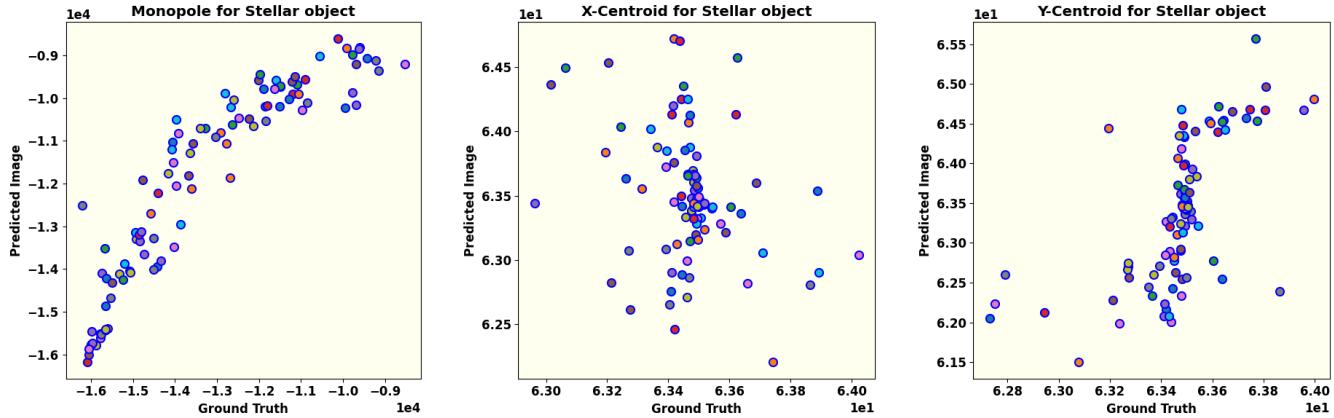
774 In this section, we discuss the results of phase retrieval  
 775 and image reconstruction by our cGAN model followed  
 776 by an analysis of these results using moments. The hy-  
 777 perparameter values that produce the best performance  
 778 in our cGAN model are tabulated in Table 1.

### 781 6.1. Predicted Image from the Trained cGAN

782 Fig. 9 demonstrates the success of the trained cGAN  
 783 model in training a model in reconstructing the images  
 784 of fast-rotating stars using their Intensity Interferometry  
 785 (II) observation . The GAN was trained on the training  
 786 datasets for 100,000 steps and subsequently tested on  
 787 various validation datasets to produce predicted images  
 788 of the stars. The training was performed on a CPU  
 789 using two nodes, each with 96 threads, and required  
 790 approximately 20 hours. In Fig. 9, four combined images  
 791 illustrate the GAN's performance in reconstructing the  
 792 stars' shape, size, and brightness distribution using II.

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

- 795 • The left panel represents the sparse II pattern ob-  
 796 tained by the simulated observation using the four  
 797 IACTs illustrated in Fig.1. As explained earlier,  
 798 this image acts as part of the input, namely its  
 799 “condition” part. This image acts as the condi-  
 800 tion that the sky-images of the star generated by  
 801 the Generator must conform to.



**Figure 10.** This set of figures shows the comparison of monopole, x-centroid, and y-centroid for ground truth and predicted images generated by trained GAN. The monopole shows an approximate linear nature, which indicate a correlation between the parameters of the ellipsoids and the flux (monopole) moment. Unclear statement: Clustering nature around the value of 63 for both x and y centroid show the spatial center of mass of the stellar objects. Discussion and rewriting needed .

that might not be apparent through visual inspection alone.

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

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y) \quad (15)$$

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. (15), the monopole of an image is calculated as:

$$M_{00} = \sum_x \sum_y I(x, y). \quad (16)$$

The left panel of Fig. 10 displays the monopole values for 100 reconstructed images. The plot reveals an approximate 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:

$$x_c = \frac{M_{10}}{M_{00}} = \frac{\sum_{x,y} x \cdot I(x, y)}{\sum_{x,y} I(x, y)} \quad (17)$$

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

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.

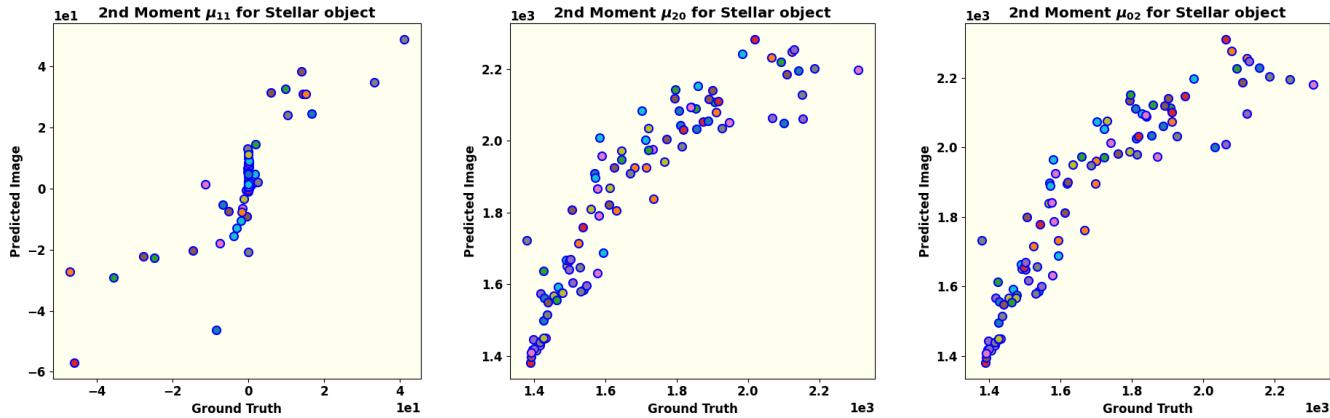
The middle and right panel of Fig. 10 compare the centroids  $(x_c, y_c)$  of 100 predicted images with their corresponding ground truths respectively. The clustering of centroids within a specific scale range across all results indicates that the reconstructed images accurately represent the spatial 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:

$$\mu_{pq} = \frac{1}{M_{00}} \sum_x \sum_y (x - x_c)^p (y - y_c)^q I(x, y). \quad (18)$$

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

Fig. 11 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



**Figure 11.** The second-order central moments from left to right ( $\mu_{11}$ ,  $\mu_{20}$ , and  $\mu_{02}$ ) provide information about the size and shape of stellar objects. These have been shown here for ground truth and predicted images generated by the trained GAN. It is very clear that these moments show the linear nature. The indicative approximate linear relation arguably points towards a close correlation between the predicted and the ground truth images.

the upcoming subsection). All three plots demonstrate an approximately 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. 12 presents all third-order moments for both the ground truth and the reconstructed image. The skewness along the  $x$  and  $y$  axes ( $\mu_{30}$  and  $\mu_{03}$ ) appears acceptable, as shown in both upper panel of Fig. 12, where a linear relationship exists between the ground truth and predicted images. However, the other higher-order moments ( $\mu_{21}$  and  $\mu_{12}$ )—particularly  $\mu_{12}$ , as depicted in both lower panel of Fig. 12—do not align as well. This indicates that further improvement is possible and should be investigated.

*This statement does not match with corresponding figure caption of Fig. 12: Discussion and result*

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

$$\theta = \frac{1}{2} \arctan \left( \frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right). \quad (19)$$

The semi-major and semi-minor axes of the stellar object are computed using the second-order central mo-

ments and are denoted as  $a$  and  $b$ , respectively.

$$\begin{aligned} a &= 2\sqrt{mp + \delta} \\ b &= 2\sqrt{mp - \delta} \end{aligned} \quad (20)$$

where,

$$mp = \frac{\mu_{20} + \mu_{02}}{2} \quad (21)$$

and

$$\delta = \frac{\sqrt{4\mu_{11}^2 + (\mu_{20} - \mu_{02})^2}}{2}. \quad (22)$$

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

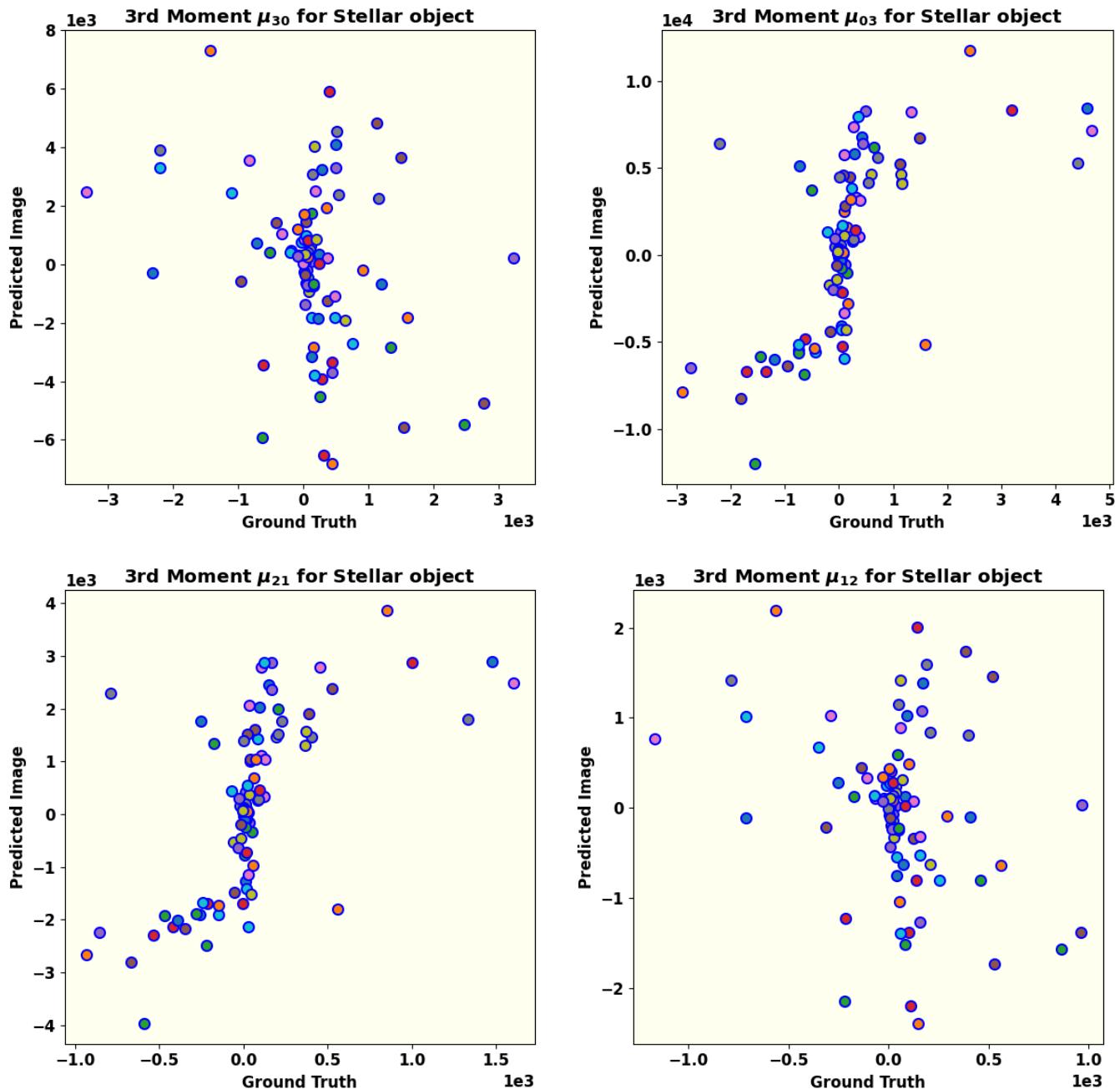
$$e = \sqrt{1 - a/b}. \quad (23)$$

Eqs. 19-23 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.

## 7. 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 a major 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). Our study demonstrates that applying a cGAN to II data



**Figure 12.** Shown here are all the third-order central moments ( $\mu_{30}$ ,  $\mu_{03}$ ,  $\mu_{21}$ , and  $\mu_{12}$ ) for ground truth and predicted images generated by the trained GAN. They represent the skewness of the brightness distributions. Here  $\mu_{03}$ , and  $\mu_{21}$  show the linear behaviour. However,  $\mu_{30}$ , and  $\mu_{12}$  show the clustering nature, which says that skewness is preserved but not perfectly. This aim can be achieved by increasing the number of telescopes on observational plane.

successfully recovers the size, shape, and brightness distribution of a fast rotating star. That the size, shape and brightness distribution of fast rotating stars can be recovered by a cGAN trained on the combined input of the sky-image of known sources along with their respective II data. In this training the sky-image acts as the real “ground truth” and the II data acts as the “condition”.  
 970  
 971  
 972  
 973  
 974  
 975  
 976

977 The Discriminator of the cGAN model is trained to efficiently distinguish between the real images and fake  
 978 (generated) images based on the “ground truth” images  
 979 and the respective II data condition. The Generator is  
 980 trained to generate progressively realistic images which  
 981 are also consistent with the condition of the II data.  
 982 The evaluation of the trained model is then carried out  
 983

984 Evaluations based on by comparison of image moments  
985 of real images and generated images. Specifically, the  
986 monopole, second, and third-order moments are compared.  
987 The results support the effectiveness of cGAN  
988 in achieving the stated objective. accurate image recon-  
989 struction from a simulation of II from a single site with  
990 four telescopes.

991 While the results of this study highlight the significant  
992 potential of machine learning, and in particular the ap-  
993 plicability of cGAN, for image reconstruction in II, sev-  
994 eral aspects require further refinement. First, an impor-  
995 tant factor in the reconstruction process is the extent of  
996 Fourier plane coverage, which depends on the number of  
997 available telescopes and the total observing time. The  
998 reasonable success of this piece of work suggests that a  
999 network of higher number of telescopes providing higher  
1000 number of baselines and greater coverage of the  $(u, v)$   
1001 plane signal, would play a critical role in projects of im-  
1002 age reconstruction of more complicated stellar systems  
1003 can be undertaken. Future work might explore different  
1004 observatory layouts to assess their impact on image re-  
1005 construction quality. Second, detector efficiencies, which  
1006 impact the signal-to-noise ratio (SNR) of actual observa-

1007 tional data, have not yet been incorporated; addressing  
1008 these factors will be crucial for more accurate SNR es-  
1009 timation. Third, exploring and comparing alternative  
1010 methods for image generation could reveal approaches  
1011 that outperform cGAN in reconstructing stellar images  
1012 with II. Fourth, experimenting with different loss func-  
1013 tions could provide additional insights into the recon-  
1014 struction quality. Although further testing is needed to  
1015 refine the GAN and enhance its robustness and reli-  
1016 ability, our findings suggest that machine learning is a  
1017 promising approach for phase reconstruction in II.

## ACKNOWLEDGEMENTS

1018  
1019 **Acknowledgment:** One of the authors (SS) gratefully  
1020 acknowledges the computing facilities and the local hos-  
1021 pitality extended to him by the Inter-University Centre  
1022 for Astronomy and Astrophysics (IUCAA), Pune, In-  
1023 dia under its Visiting Associate Programme during the  
1024 preparation of this manuscript.

## REFERENCES

- 1025 Abadi et al., M. 2016, arXiv preprint arXiv:1603.04467  
1026 Abe et al., S. 2024, MNRAS, 529, 4387  
1027 Acciari et al., V. A. 2020, MNRAS, 491, 1540  
1028 Acharyya et al., A. 2024, The Astrophysical Journal, 966,  
1029 28  
1030 Aleksić et al., J. 2016, Astroparticle Physics, 72, 76  
1031 Archer, A., Aufdenberg, J. P., Bangale, P., et al. 2025,  
1032 arXiv e-prints, arXiv:2506.15027,  
1033 doi: [10.48550/arXiv.2506.15027](https://doi.org/10.48550/arXiv.2506.15027)  
1034 Baumgartner et al., S. 2020, MNRAS, 498, 4577  
1035 Coccolini et al., D. 2021, arXiv preprint arXiv:2122.11578  
1036 Domiciano de Souza, A., Kervella, P., Jankov, S., et al.  
1037 2003, Astronomy and Astrophysics, 407, L47,  
1038 doi: [10.1051/0004-6361:20030868](https://doi.org/10.1051/0004-6361:20030868)  
1039 Domiciano de Souza, A., Kervella, P., Jankov, S., et al.  
1040 2005, Astronomy and Astrophysics, 442, 567,  
1041 doi: [10.1051/0004-6361:20042476](https://doi.org/10.1051/0004-6361:20042476)  
1042 Dravins et al., D. 2013, Astroparticle Physics, 43, 331,  
1043 doi: [10.1016/j.astropartphys.2012.04.017](https://doi.org/10.1016/j.astropartphys.2012.04.017)  
1044 Fienup, J. 1982, Applied Optics, 21, 2758  
1045 Gamo, H. 1963, Journal of Applied Physics, 34, 875  
1046 Gerchberg, R. W. 1972, Optik, 35, 237  
1047 Glauher, R. J. 1963, Physical Review, 130, 2529  
1048 Goldberger et al., M. L. 1963, Physical Review, 132, 2764  
1049 Goodfellow et al., I. 2014, Advances in neural information  
1050 processing systems, 27  
1051 Hanbury Brown, R., & Twiss, R. Q. 1956, Nature, 177, 27  
1052 Hanbury Brown et al., R. 1957, Proceedings of the Royal  
1053 Society of London. Series A. Mathematical and Physical  
1054 Sciences, 242, 300  
1055 Hanbury Brown et al., R. 1958, Proceedings of the Royal  
1056 Society of London. Series A. Mathematical and Physical  
1057 Sciences, 243, 291  
1058 Hanbury Brown et al., R. 1974, Monthly Notices of the  
1059 Royal Astronomical Society, 167, 121  
1060 Haubois et al., X. 2009, Astronomy & Astrophysics, 508,  
1061 923, doi: [10.1051/0004-6361/200912927](https://doi.org/10.1051/0004-6361/200912927)  
1062 Holmes et al., R. 2010, in Adaptive Coded Aperture  
1063 Imaging, Non-Imaging, and Unconventional Imaging  
1064 Sensor Systems II, Vol. 7818, SPIE, 175–185  
1065 Hu, M.-K. 1962, IRE transactions on information theory, 8,  
1066 179  
1067 Hunter, J. D. 2007, Computing in Science & Engineering, 9,  
1068 90, doi: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55)  
1069 Isola et al., P. 2017, in Proceedings of the IEEE conference  
1070 on computer vision and pattern recognition, 1125–1134  
1071 Kirisits et al., C. 2024, arXiv preprint arXiv:2406.14143v2  
1072 Le Bohec, S.; Holder, J. 2006, The Astrophysical Journal,  
1073 649, 399

- 1074 Li et al., X. 2014, in Imaging Spectroscopy, Telescopes and  
1075 Large Optics, Vol. 9298, SPIE, 92981G1–G7
- 1076 Liu, L.-C., Wu, C., Li, W., et al. 2024, arXiv preprint,  
1077 arXiv:2404.15685, doi: [10.48550/arXiv.2404.15685](https://doi.org/10.48550/arXiv.2404.15685)
- 1078 Liu et al., L.-C. 2025, Physical Review Letters, 134, 180201,  
1079 doi: [10.1103/PhysRevLett.134.180201](https://doi.org/10.1103/PhysRevLett.134.180201)
- 1080 Lucy, L. B. 1967, Zeitschrift für Astrophysik, Vol. 65, p. 89,  
1081 65, 89
- 1082 Maeder, A. 1999, Astronomy & Astrophysics, 347, 185
- 1083 Martinez, A. O., Baron, F., Monnier, J. D., Roettenbacher,  
1084 R. M., & Parks, J. R. 2021, The Astrophysical Journal,  
1085 916, 60, doi: [10.3847/1538-4357/ac06a5](https://doi.org/10.3847/1538-4357/ac06a5)
- 1086 McAlister et al., H. A. 2005, The Astrophysical Journal,  
1087 628, 439
- 1088 Mirza et al., M. 2014, arXiv preprint arXiv:1411.1784
- 1089 Monnier, J. D., Zhao, M., Pedretti, E., et al. 2007, Science,  
1090 317, 342, doi: [10.1126/science.1143205](https://doi.org/10.1126/science.1143205)
- 1091 Murphy, K. P. 2022, Probabilistic machine learning: an  
1092 introduction (MIT press)
- 1093 Mustafa et al., M. 2019, Computational Astrophysics and  
1094 Cosmology, 6, 1
- 1095 Norris et al., R. P. 2021, The Astrophysical Journal, 919,  
1096 124, doi: [10.3847/1538-4357/ac0c7e](https://doi.org/10.3847/1538-4357/ac0c7e)
- 1097 Nuñez et al., P. D. 2010, in Optical and Infrared  
1098 Interferometry II, Vol. 7734, SPIE, 458–467
- 1099 Nuñez et al., P. D. 2012a, Monthly Notices of the Royal  
1100 Astronomical Society, 419, 172
- 1101 Nuñez et al., P. D. 2012b, Monthly Notices of the Royal  
1102 Astronomical Society, 424, 1006
- 1103 Pedretti, E., Monnier, J. D., ten Brummelaar, T., &  
1104 Thureau, N. D. 2009, New Astronomy Reviews, 53, 353,  
1105 doi: [10.1016/j.newar.2010.07.005](https://doi.org/10.1016/j.newar.2010.07.005)
- 1106 Prince, S. J. 2023, Understanding deep learning (MIT  
1107 press)
- 1108 Rai, K. N., Basak, S., Sarangi, S., & Saha, P. 2025,  
1109 Resonance, 30, 45, doi: [10.1007/s12045-025-1729-x](https://doi.org/10.1007/s12045-025-1729-x)
- 1110 Rai et al., K. N. 2021, Monthly Notices of the Royal  
1111 Astronomical Society, 507, 2813,  
1112 doi: [10.1093/mnras/stab2391](https://doi.org/10.1093/mnras/stab2391)
- 1113 Rai et al., K. N. 2022, Monthly Notices of the Royal  
1114 Astronomical Society, 516, 2864,  
1115 doi: [10.1093/mnras/stac2433](https://doi.org/10.1093/mnras/stac2433)
- 1116 Ronneberger et al., O. 2015, in Medical image computing  
1117 and computer-assisted intervention—MICCAI 2015: 18th  
1118 international conference, Munich, Germany, October 5–9,  
1119 2015, proceedings, part III 18, Springer, 234–241
- 1120 Sato et al., T. 1978, Applied optics, 17, 2047
- 1121 Sato et al., T. 1979, Applied Optics, 18, 485
- 1122 Sato et al., T. 1981, Applied Optics, 20, 2055
- 1123 Schawinski et al., K. 2017, MNRAS, 467, L110
- 1124 Teague, M. R. 1980, Josa, 70, 920
- 1125 Teague, M. R. 1983, Journal of Optical Society of  
1126 America, 73, 1434
- 1127 van Belle, G. T., Ciardi, D. R., Thompson, R. R., Akeson,  
1128 R. L., & Lada, E. A. 2001, The Astrophysical Journal,  
1129 559, 1155, doi: [10.1086/322340](https://doi.org/10.1086/322340)
- 1130 Virtanen et al., P. 2020, Nature methods, 17, 261
- 1131 Vogel et al., N. 2025, MNRAS, 537, 2334
- 1132 Von Zeipel, H. 1924, Monthly Notices of the Royal  
1133 Astronomical Society, Vol. 84, p. 665–683, 84, 665
- 1134 Zhang et al., J. 2020, Opt. Lett., 45, 3649
- 1135 Zhao, M., Monnier, J. D., Pedretti, E., et al. 2009, The  
1136 Astrophysical Journal, 701, 209,  
1137 doi: [10.1088/0004-637X/701/1/209](https://doi.org/10.1088/0004-637X/701/1/209)