

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

45 reconstruction of a fast rotator using its simu-  
46 lated II observations.

47 The foundational basis of II stems from the pi-  
48 oneering experiments and theoretical investiga-  
49 tions carried out initially by Hanbury Brown and  
50 Twiss (HBT) (R. Hanbury Brown & R. Q. Twiss  
51 1956; R. Hanbury Brown et al. 1957, 1958) and,  
52 later, by (R. J. Glauber 1963) and others. The  
53 correlation between photons (called the , widely  
54 referred to as the “HBT effect”) measured by a  
55 pair of photon detectors in two partially coher-  
56 ent beams of light was reported in 1956 R. Han-  
57 bury Brown & R. Q. Twiss (1956). K. N. Rai  
58 et al. (2025) present a recent variant of this ex-  
59 periment, carried out with pseudo-thermal light.  
60 The HBT Effect and the related theoretical in-  
61 vestigations laid the foundation for the modern  
62 field of Quantum Optics.

63 Hanbury Brown and his collaborators led the creation  
64 and installation of the historic II facility at Narrabri,  
65 Australia, and reported the measurement of angular di-  
66 ameters of 32 stars and a few multiple star-systems (R.  
67 Hanbury Brown et al. 1974). Soon after this work, how-  
68 ever, II observations of stars was stalled for over four  
69 decades due to the limits of the then available photon  
70 detectors and data processing equipment. With grad-

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

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

Furthermore, given that achieving a high signal-to-noise ratio (SNR) is often challenging in astronomical datasets, it is immensely beneficial if additional data can be generated using the available information from the observed sky density distribution and ground-based observations (II data, in our case) of the sources under investigation. Generative Adversarial Networks (GANs), introduced by I. Goodfellow et al. (2014), have been successful in such data augmentation tasks. Conditional

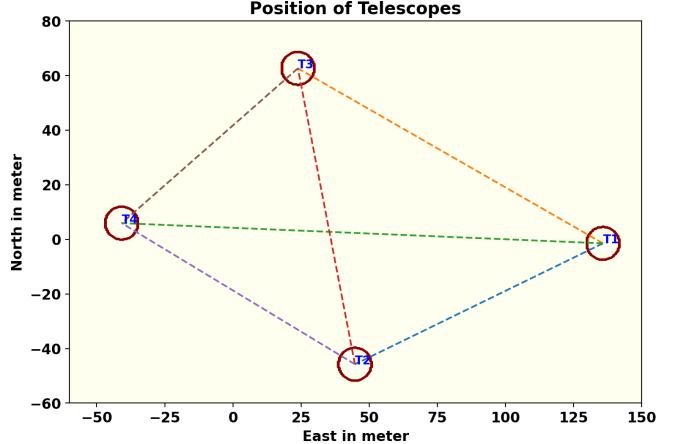


Figure 1. A schematic representation of a hypothetical observation facility with an array of four Cherenkov Telescopes (IACTs) whose relative positions approximately mimic those at VERITAS. This array is used in simulating the II observation of the fast rotators. Each of the telescopes has a diameter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

GAN (cGAN) architectures, proposed by M. Mirza et al. (2014) and applied to a wide variety of datasets by P. Isola et al. (2017), leverage additional information about the images in the training datasets and have demonstrated remarkable robustness in image recovery across diverse data types.

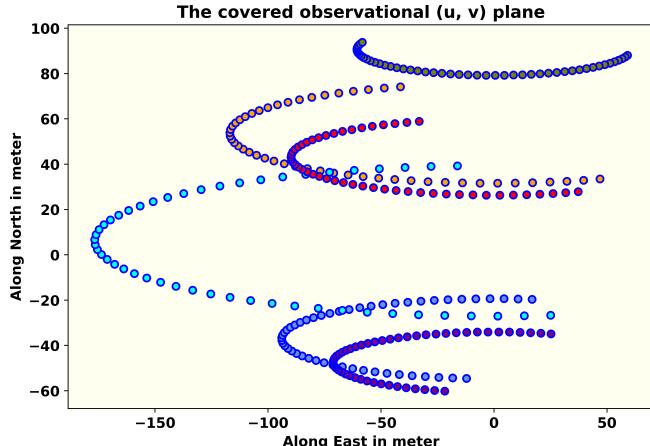
In the astrophysical context, K. Schawinski et al. (2017) employed a GAN model to recover features, such as spiral arms, central bulges, and disk structures of galaxies, from noise-affected images. M. Mustafa et al. (2019) developed and customized a Deep Convolutional GAN, dubbed “CosmoGAN”, capable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

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

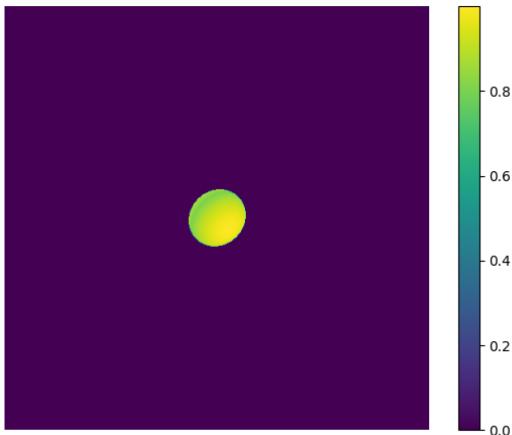
This section presents a brief conceptual overview of how an array of telescopes is used to perform II observations, and explains the Signal-to-Noise Ratio (SNR) from these measurements.

#### 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 simultane-



**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.**



**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:

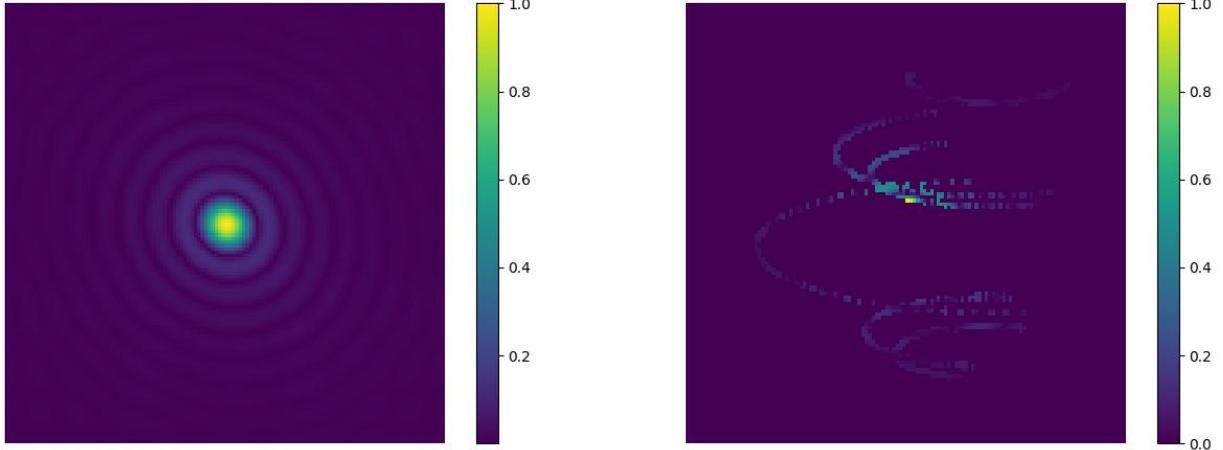
$$c = g^{(2)} - 1 = \frac{\Delta f}{\Delta \nu} |V_{12}|^2 \quad (3)$$

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 photomultiplier 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



**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.

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  $b$ .

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

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.

### 365 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

### 412 413 problem along with mitigation of loss of phase information in II.

414 II counts the photons arriving at the telescopes from  
 415 the stellar object. The correlation of these photon  
 416 arrivals at the telescopes yields the squared visibility  
 417 Eq. (5), as explained in subsection 3.1. The left panel  
 418 of fig. 4 shows the signal from the source shown in Fig. 3  
 419 using II, displayed on linear scales. A point to note  
 420 here is that this figure represents the signal from the  
 421 source that would be recorded by an infinite number of  
 422 baselines provided by an infinite number of telescopes on  
 423 the interferometric plane. In practice, only a small part  
 424 of this information is available (as seen in right panel  
 425 of Fig. 4), because one has a finite number of baselines  
 426 corresponding to the finite number  $N_T$  of telescopes at  
 427 our disposal and a limited observation schedule. We  
 428 have simulated the II observation of the fictitious star  
 429 by a hypothetical observation facility having an  
 430 array of four IACTs with their relative positions  
 431 approximating those at VERITAS (correlated with  
 432 baselines as seen in Fig. 1) over one night. Using this  
 433 modest amount of signal from one night’s observation,  
 434 we have trained a cGAN to construct the image of the  
 435 source.

### 437 4. GENERATIVE ADVERSARIAL NETWORKS

438 Generative Adversarial Networks (GANs) were in-  
 439 troduced by I. Goodfellow et al. (2014). A GAN  
 440 model involves two competing deep neural net-  
 441 work models, referred to as the Generator and the  
 442 Discriminator. These two networks engage in a  
 443 zero-sum “Minimax” game, as in Game Theory.  
 444 Given a real data set  $\{x_i\}$  (for example, a set of  
 445 real images) drawn from some unknown distri-  
 446 bution  $P_{\text{data}}(x)$  generated by some unknown or  
 447 ill-understood process, the objective here is to  
 448 generate a new set (of images) whose probability  
 449 distribution should match  $P_{\text{data}}(x)$  as closely as  
 450 possible. During the training of the two models,  
 451 the Generator samples a latent variable  $z$  from a  
 452 known prior distribution  $P_z(z)$  (e.g., the Normal  
 453 Distribution) and produces a synthetic sample  
 454  $G(z)$  to start with and, subsequently, based on  
 455 updates received from the Discriminator as its  
 456 training progresses. The Discriminator, being a  
 457 probabilistic binary classifier, receives either a  
 458 real data sample  $x$  or a generated sample  $G(z)$ ,  
 459 and outputs a probability that the input is real.  
 460 The Discriminator aims to maximise its classi-  
 461 fication accuracy, while the Generator aims to  
 462 fool it by trying to minimize it. The training of  
 463 these two networks proceeds alternately leading



**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.

464 to the optimization of the adversarial loss func-  
465 tion  $L(D, G)$  given by

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

466  
467 The two neural networks  $G(z) \equiv G(z; \theta_G)$  and  
468  $D(x) \equiv D(x; \theta_D)$  are parameterized by  $\theta_i$  with  
469  $i = G$  or  $D$  respectively. During its training,  
470 the Generator generates the differentiable func-  
471 tion  $x_{\text{gen}} \equiv G(z)$  that maps  $z$  to the data space  
472  $x$ . Through such maps the Generator gener-  
473 ates its own distribution  $p_G(x_{\text{gen}})$  and through the  
474 training episodes, specifically by iteratively up-  
475 dating its parameters  $\theta_G$ , aligns this distribution  
476 with the distribution of real data  $p_{\text{data}}(x)$ . The  
477 training data set provided to the Discriminator  
478 is constructed by mixing real data points  $x$  and  
479 generated data points  $x_{\text{gen}}$  in equal ratio. The  
480 Discriminator generates the function  $D(x)$  that  
481 represents the probability that  $x$  originates from  
482 real data. Eq.(9) implies that training of the  
483 GAN model moves towards maximization of the  
484 expectation of  $D(x)$ . Through this process, both

485 the parameters  $\theta_G$  and  $\theta_D$  are optimized such that  
486  $p_G(x_{\text{gen}})$  gets maximally aligned with  $p_{\text{data}}(x)$ .

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

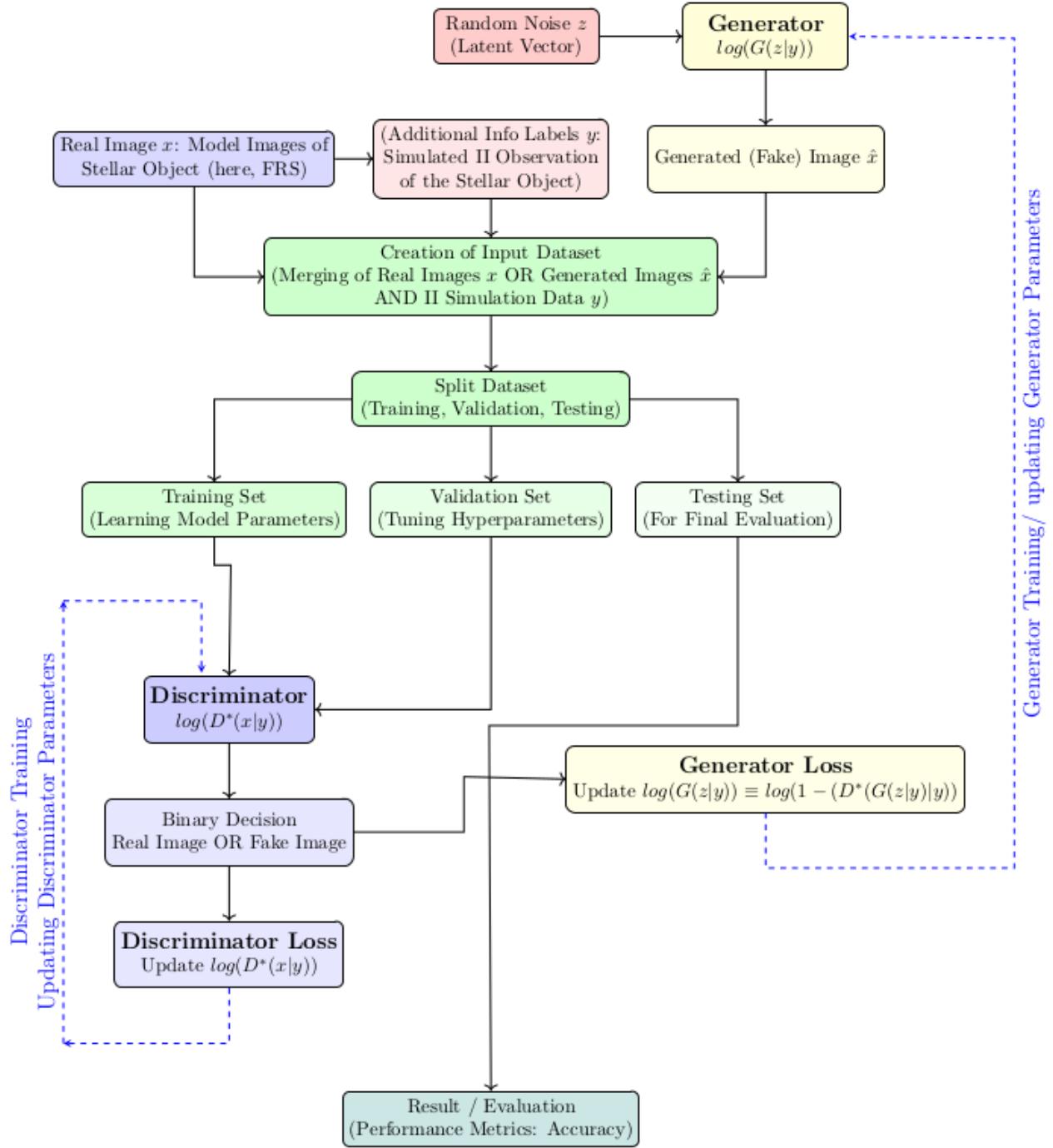
$$\begin{aligned} \max_D V(D, G) &= \mathbb{E}_{x \sim p_{\text{data}}} [\log D_G^*(x)] \\ &\quad + \mathbb{E}_{x \sim p_G} [\log (1 - D_G^*(x))] \end{aligned} \quad (10)$$

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

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

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

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



**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.

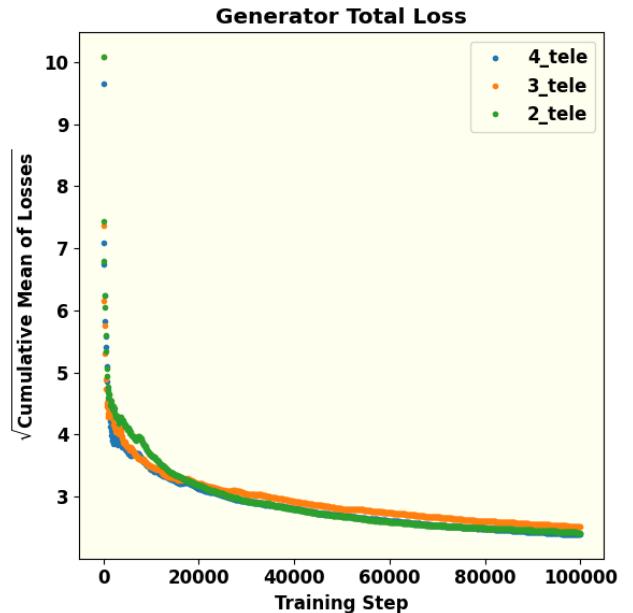
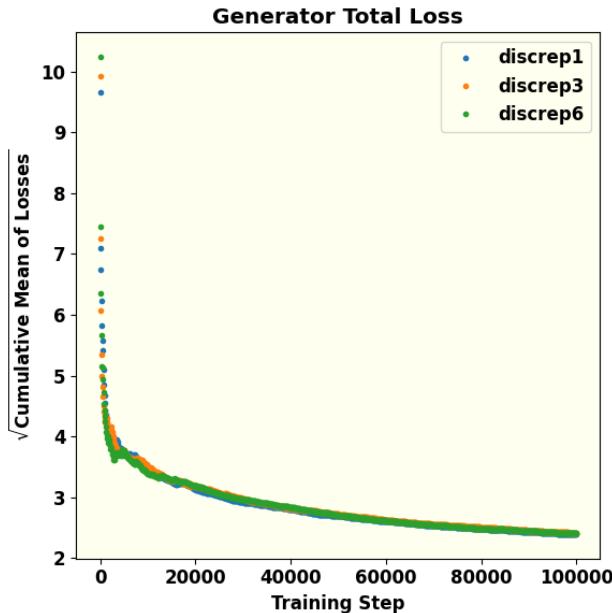
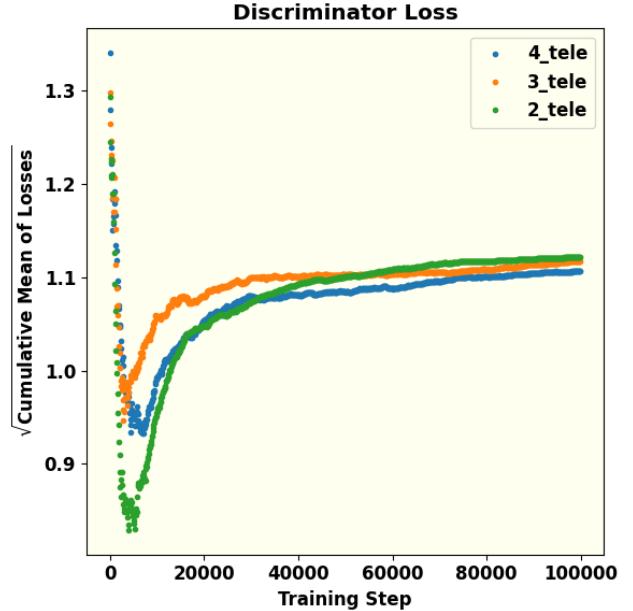
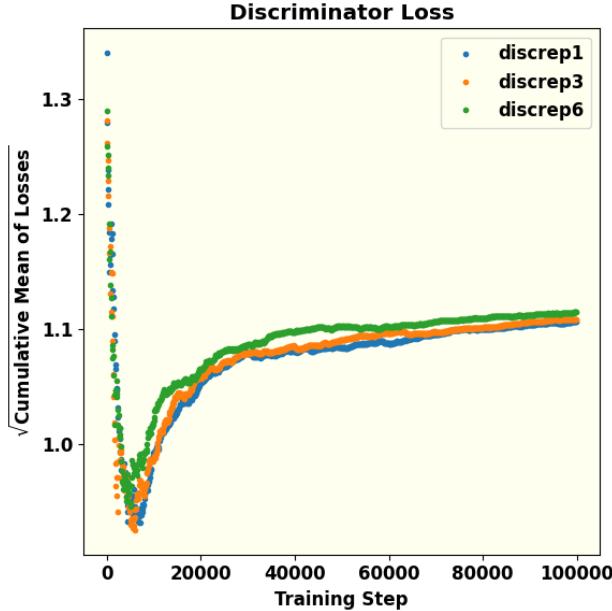


Figure 7. These figures show the effect of the ratio of episodes of Discriminator training per every episode of Generator training. This hyperparameter is termed Discriminator repetition or (discrep) in the figures. The square root of the cumulative mean of the losses are plotted against the training 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.

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

$$507 \quad V(D, G) = \mathbb{E}_{x, y \sim p_{\text{data}}(x)} [\log D(x|y)] \\ + \mathbb{E}_{z, y \sim p_z(z)} [\log (1 - D(G(z|y)|y))] \quad (12)$$

508 The conditional variable  $y$  in a cGAN can be  
 509 various additional information including images,  
 510 labels or text contextual to “ground truth” real  
 511 data  $x$ . Among various types of cGANs, Pix2Pix  
 512 GAN with its image-to-image translation de-  
 513 sign is suitable for the task at hand. In our  
 514 work, we specifically use this conditional vari-  
 515 able by choosing  $y$  to represent the ground-  
 516 based intensity-interferometry (II) observation  
 517 patterns: the Generator is trained to produce  
 518 stellar surface images that not only look realistic  
 519 but also conform to the measured II correlations,  
 520 while the Discriminator judges realism *and* con-  
 521 sistency with the II data.

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

$$529 \quad L_1(G) = \mathbb{E}_{x \sim p_{\text{data}}, z \sim p_z(z)} [\|x - G(z|y)\|_1]. \quad (13)$$

530 The total adversarial loss function, along with  
 531 the  $L_1$  loss modulated by a hyperparameter  $\lambda$   
 532 then becomes

$$533 \quad L_{\text{tot}} = \arg \min_G \max_D V(D, G) + \lambda \cdot L_1(G). \quad (14)$$

534 This type of network has demonstrated remarkable ro-  
 535 bustness across a variety of applications. For example, it  
 536 can generate colored images from grayscale inputs based  
 537 on architectural labels, transform images from day to  
 538 night, and even predict maps from satellite data. A more  
 539 extensive list of applications is provided in P. Isola et al.  
 540 (2017).

#### 541 4.1. Generator

542 As discussed above, in a GAN the Generator, a deep  
 543 neural network in itself, is responsible for producing  
 544 synthetic data, in this case, images that resemble those  
 545 of a fast-rotating star. In this work, the Generator  
 546 is implemented as a U-Net convolutional network (O.  
 547 Ronneberger et al. 2015). The U-Net consists of a  
 548 symmetric encoder–decoder structure forming a  
 549 characteristic “U” shape along with skip connec-  
 550 tions: the left (contracting or down-sampling)

551 path, the right (expanding or up-sampling path)  
 552 and the connecting (bottle-neck) path. The  
 553 left down-sampling path repeatedly applies  $3 \times 3$   
 554 convolutions (padded to preserve spatial dimen-  
 555 sions) followed by LeakyReLU activations and  
 556 strided convolution (with stride of 2), by progres-  
 557 sively halving the spatial resolution while dou-  
 558 bling the channel depth (typically  $64 \rightarrow 128 \rightarrow$   
 559  $256 \rightarrow 512 \rightarrow 1024$ ). At the bottleneck, high-level  
 560 features are processed without further spatial re-  
 561 duction. The right up-sampling path mirrors  
 562 this process using  $2 \times 2$  transposed convolutions  
 563 (stride 2) for up-sampling, halving the channel  
 564 count at each level and using ReLU as the acti-  
 565 vation function for all its layers except the out-  
 566 put layer. Additionally, a dropout layer is introduced  
 567 at the beginning of the upsampling phase to mitigate  
 568 overfitting of the Generator model (P. Isola et al. 2017).  
 569 Critically, long skip connections concatenate fea-  
 570 ture maps from each encoder (down-sampling)  
 571 level to the corresponding decoder (up-sampling  
 572 level), directly injecting high-resolution details  
 573 into the reconstruction process. This enables  
 574 the network to “remember where everything  
 575 is” while the deep bottleneck still provides the  
 576 large receptive field needed to think globally  
 577 about image structure and semantics. Besides  
 578 the strided convolutions, modern variants of  
 579 U-Nets used in state-of-the-art GANs incorpo-  
 580 rate residual blocks within resolution levels, and  
 581 frequently add spectral normalization and self-  
 582 attention at the bottleneck for improved train-  
 583 ing stability and long-range dependency mod-  
 584 elling. These architectural choices allow our U-  
 585 Net generator to simultaneously recover sharp,  
 586 high-frequency details (stellar limb edges, limb-  
 587 darkening profiles, gravity darkening, and rapid-  
 588 rotation-induced oblateness) and enforce global  
 589 coherence (overall disk morphology and physi-  
 590 cal consistency with the observed interferometric  
 591 visibilities)—making it ideally suited for high-  
 592 fidelity sky-image reconstruction of fast-rotating  
 593 stars from sparse ground-based intensity inter-  
 594 ferometry observations.

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

$$598 \quad L_G = -\mathbb{E}_{z \sim p_z(z)} [\log D(G(z|y) | y)] \\ + \lambda \mathbb{E}_{x \sim p_{\text{data}}, z \sim p_z(z)} [\|x - G(z|y)\|_1]. \quad (15)$$

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

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

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

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

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

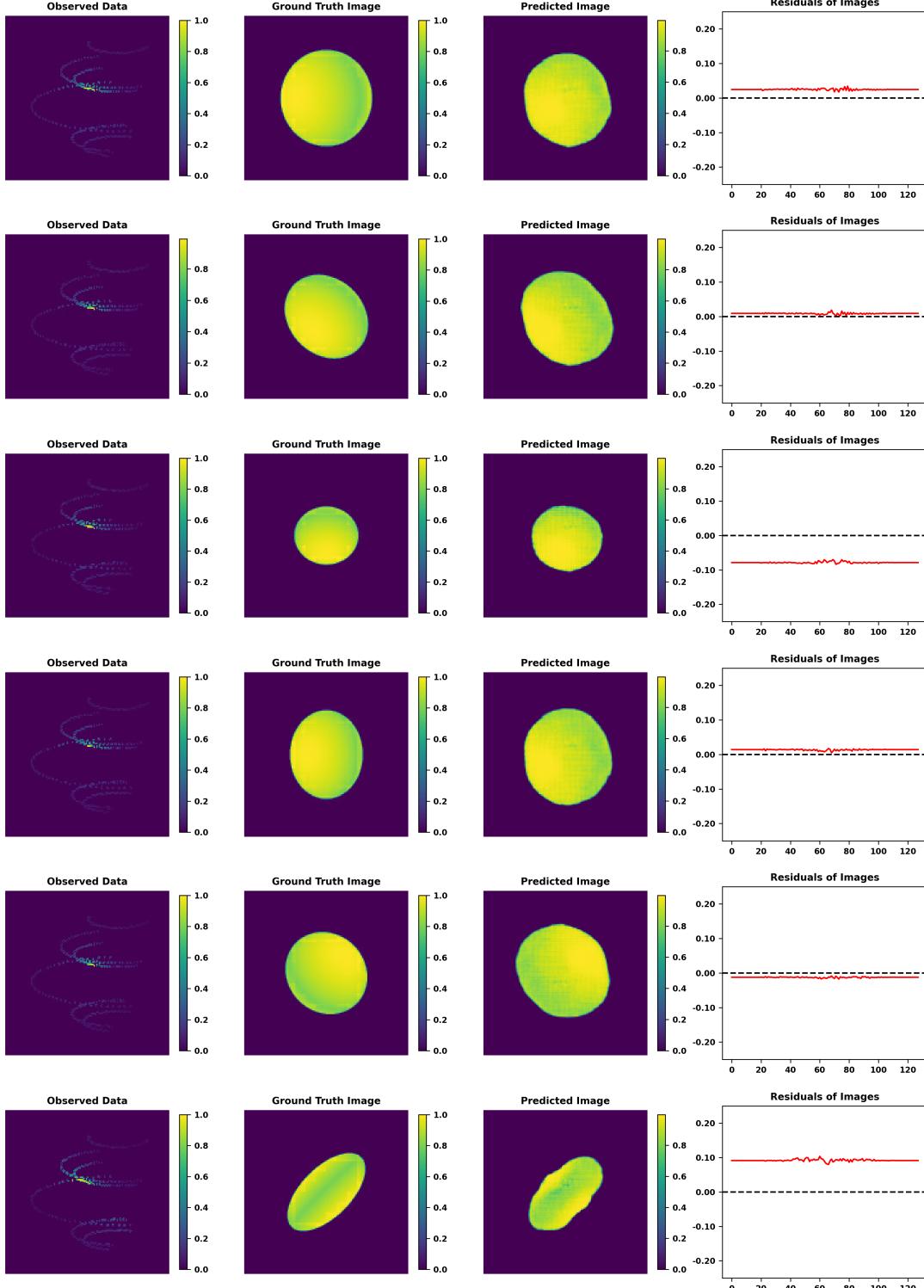
- The discriminator  $D$  is updated by minimizing  $L_D$ , keeping  $G$  fixed (so that  $D$  becomes better at distinguishing real vs generated images).
- The generator  $G$  is updated by minimizing  $L_G$ , keeping  $D$  fixed, thus pushing  $G$  to generate images that are (a) judged “real” by 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

### 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 generate synthetic images of rapidly rotating stars by modelling them as oblate spheroids with varying radii and oblateness parameters between 0.5 and 1.



**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.

696 To incorporate the effect of gravity darkening,  
 697 we also consider different viewing angles and assume a  
 698 linear dependence on the declination angle of each  
 699 point on the stellar surface. The traced ellipses re-  
 700 sult from integrating over the source's hour angle.

701 Next, Salt and Pepper noise is introduced into these  
 702 images; usually this is done at the rate of 0.5%  
 703 of the number of pixels in the image. Then, the  
 704 images are resized and their mean is subtracted. A  
 705 two-dimensional Fast Fourier Transform, along with a  
 706 Fourier shift, is applied, yielding a complex number for  
 707 each pixel. Since II does not measure phase, the abso-  
 708 lute value is calculated.

709 Next, sampling of the interferometric plane is  
 710 introduced via pixel-wise multiplication of the  
 711 absolute-valued Fourier-transformed image (left  
 712 panel of Fig. 4) and the baseline tracks of (Fig. 2)  
 713 generated by the four IACTs due to Earth rota-  
 714 tion during the II observation. The result is a map  
 715 in the Fourier plane featuring several ellipses, which is  
 716 also referred to as the sparse sampling map (right panel  
 717 of fig. 4). This map represents the sparse sampling of  
 718 the signal space corresponding to the source (Fig. 3) ob-  
 719 served with four telescopes (Fig. 1). In essence, the  
 720 right panel of fig. 4 represents the result of simu-  
 721 lated II observation of the fictitious fast-rotator  
 722 illustrated in fig. 3.

723 Finally, we normalize the pixel values and convert  
 724 them to 8-bit integers, producing an image that encodes  
 725 the sparsely sampled, phase-free visibility measured by  
 726 II. The image of the corresponding simulated (ficti-  
 727 tious) star, which serves as the “ground truth”  
 728 is processed identically to avoid any bias. These  
 729 two images are merged side-by-side into a single  
 730 image (as shown in Fig. 5). The dataset thus cre-  
 731 ated is split into three parts in the proportions  
 732 of 80% for training, 10% for validation, and 10%  
 733 as the test set. This partition of the full dataset  
 734 is indicated in the Fig. 6. In the following, we refer  
 735 to these three parts as the Training Set, the  
 736 Validation Set and the Testing Set respectively.

## 737 5.2. GAN Architecture

738 The GAN architecture used in this work is a Pix2Pix  
 739 cGAN, which uses an image-to-image translation  
 740 strategy with both the ground truth and the con-  
 741 dition being images. Originally introduced by  
 742 Isola et al. P. Isola et al. (2017), this architec-  
 743 ture is specifically suitable for image processing  
 744 and reconstruction objectives. For instance, the

745 TensorFlow tutorials<sup>6</sup> demonstrate its application to a  
 746 dataset of architectural facades. This architecture has  
 747 been adapted for the phase retrieval problem at hand  
 748 here. The network is implemented using the Tensor-  
 749 Flow library (M. Abadi et al. 2016), calculations are  
 750 performed with scipy (P. Virtanen et al. 2020), and plots  
 751 are generated with matplotlib (J. D. Hunter 2007).

### 752 5.3. Hyperparameter Tuning

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

757 The learning rate ( $lr$ ) of the optimizer determines how  
 758 much the model updates its parameters with each it-  
 759 eration. A learning rate that is too small may lead  
 760 to underfitting, while one that is too large can render  
 761 the model unstable. Therefore, selecting an appropriate  
 762 learning rate is crucial (K. P. Murphy 2022). A canonical  
 763 choice in Pix2Pix and other GAN models is  
 764  $lr = 2 \times 10^{-4}$ . In our case too, we found this choice  
 765 to be appropriate.

766 The kernel size refers to the dimensions of the con-  
 767 volutional kernel used in the network, determining how  
 768 many pixels are combined to produce a new pixel. A  
 769 larger kernel size can capture features spanning several  
 770 pixels, but it may also incorporate unrelated features.  
 771 Small kernel sizes are preferable in cases where  
 772 target images finer details or high spatial fre-  
 773 quency features. Since the “ground truth” tar-  
 774 get images in our case have longer scale gravity  
 775 darkening features, we have opted for the more  
 776 canonical choice of kernel size being 5.

777 The amount of noise is controlled by two parameters,  
 778 “alpha” and “beta”, which indicate the percentage of  
 779 pixels altered to either white or black, hence the term  
 780 Salt and Pepper noise. Here, “alpha” is applied to the  
 781 real image, while “beta” is applied to the generated im-  
 782 age. Different ratios (“alpha/beta”) can lead to varying  
 783 model performance; however, our results indicate that  
 784 distinct noise rates do not significantly affect the loss  
 785 functions.

786 The batch size defines the number of images processed  
 787 simultaneously by the network. Smaller batch sizes have  
 788 been observed to improve generalization (S. J. Prince  
 789 2023). However, because a larger batch size significantly  
 790 increases training time, a batch size of 1 is used. Be-  
 791 sides, in Pix2Pix cGAN implementations found  
 792 in literature this choice is found to be often pre-  
 793 ferred.

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

794     The buffer size of a Pix2Pix GAN refers to a  
 795     small memory of image pool of previously gen-  
 796     erated images. They are occasionally fed to the  
 797     Discriminator in place of the freshly generated  
 798     ones. This strategy of mixing old and new fakes  
 799     mitigates the risk of mode collapse wherein the  
 800     Discriminator tends to map all or large num-  
 801     ber of generated images to only one or a few  
 802     real “ground truth” image(s). We have chosen  
 803     a fairly large buffer size (=1400) to protect the  
 804     model against mode collapse.

805     In the training of GANs, one often-followed strategy  
 806     to potentially boost performance is to give the Discrimi-  
 807     nator an advantage by increasing its number of train-  
 808     ing epochs before returning to the Generator’s training.  
 809     This hyperparameter is referred to as the Discriminator  
 810     repetition (as seen in Table 1). While this can lower the  
 811     Discriminator loss, as shown in Fig. 7, it also increases  
 812     training time. In training our model, we did not no-  
 813     tice any significant advantage derived from this strategy.  
 814     Since the generated images did not noticeably improve  
 815     with additional Discriminator training, we adopted the  
 816     strategy of training both the networks with equal pref-  
 817     erence (Discriminator repetition = 1).

818     One domain specific hyperparameter of the  
 819     cGAN model presented here is the Number of  
 820     Telescopes  $N_T$ . The degree of sparse sampling  
 821     of the intensity interferometric (II) image plane  
 822     can be varied to provide the model with access to  
 823     more number of active (non-zero) pixels. Point  
 824     to note here, is that the coverage of the Fourier  
 825     interferometric plane (number of active pixels  
 826     in the II image) scales with available number  
 827     of baselines, and the latter scales quadratically  
 828     with the number of telescopes. Fig. 8 shows the  
 829     loss functions for different numbers of telescopes. There  
 830     is a significant disparity in performance of the model: at  
 831     the “Nash point”, both the loss functions are minimized  
 832     for the case of four telescopes. Overall, the degree of  
 833     sparse sampling appears to have the most pronounced  
 834     effect of all the hyperparameters.

835     The hyperparameter Output Channels refers  
 836     to the number of channels in the generator out-  
 837     put (e.g., 1 for grayscale, 3 for RGB). It is worth-  
 838     while to recall that the cGAN model constructed  
 839     in this work is trained on “ground truth” target  
 840     images and the simulated II data, both in grey  
 841     scales as seen in Fig.5. It is natural that the  
 842     output of this model will be in grey scales only.  
 843     Therefore the value of this hyperparameter is set  
 844     to 1. The choice of hyperparameter  $\lambda$  has been  
 845     commented upon earlier.

846     An optimized set of hyperparameters is se-  
 847     lected through an iterative process of training  
 848     and validation. For a tentatively chosen set of  
 849     the hyperparameters, the model is trained using  
 850     the Training Set until the both the Discrimina-  
 851     tor and the Generator loss functions are mini-  
 852     mized. This model, thus trained along with its  
 853     model parameters, is then passed through val-  
 854     idation using the Validation Set. This cycle of  
 855     training and validation is iterated till an opti-  
 856     mal set of hyperparameters is arrived at. During  
 857     each epoch of training of the model, both the  
 858     Discriminator and the Generator networks are  
 859     trained for 100,000 steps. Plots of the “Discrimi-  
 860     nator Loss” function and the “Generator To-  
 861     tal Loss” function presented in Fig.7 and Fig.8  
 862     represent the results of this training for two of  
 863     the hyperparameters, namely, the Discriminator  
 864     repetition (“discrep”, in short) and the Number  
 865     of telescopes, respectively. Obviously, the most  
 866     compute-intensive part of this process is that of  
 867     the training of the Model. The results of training  
 868     and validation presented in this work were car-  
 869     ried out on a CPU using two nodes, each with  
 870     48 threads and the entire process of training and  
 871     validation required approximately 20 hours on  
 872     the machine employed for this work. This it-  
 873     erative process of training and validation of the  
 874     Model is represented schematically in the Fig.6.  
 875     The chosen optimal set of hyperparameters is  
 876     presented in the table Table-1. This optimized

**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 Telescope	4
Output Channels	1
Lambda	100

877  
 878     and trained Model is then subjected to testing  
 879     and evaluation using the Testing Set. The results  
 880     of this testing and evaluation is presented in the  
 881     following section.

882     The Pix2Pix cGAN architecture along with  
 883     the choice of the values of the hyperparame-

ters 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° would not change the data. That is, each predicted image contains an arbitrary choice among two possible orientations, differing by 180°.
- 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 distribution, 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:

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

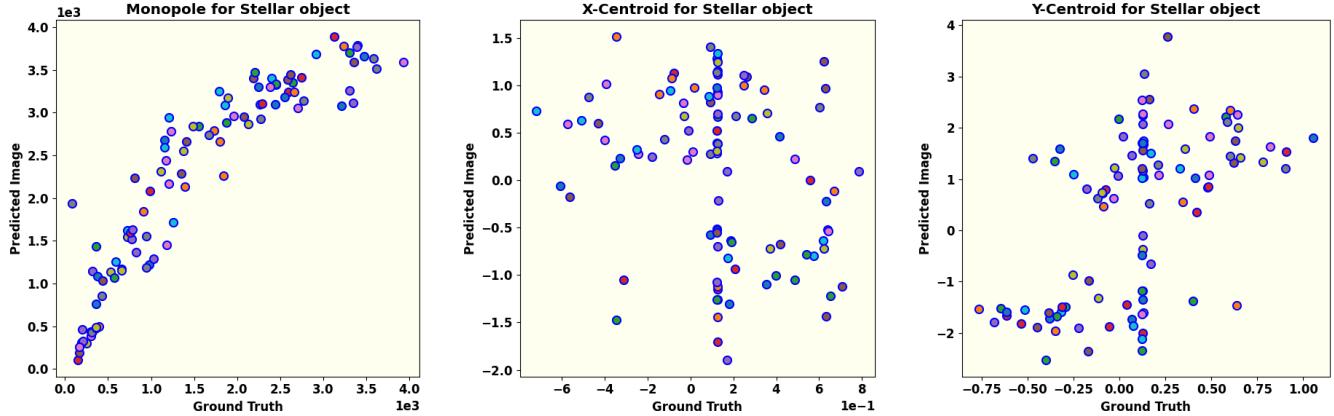


Figure 10. This set of figures shows comparisons of the monopole, and the coordinates of the centroids ( $x_c, y_c$ ) of the pixel brightness distribution of the ground truth images and the corresponding images predicted (generated) by the trained model. The monopole moment of the images represents the overall pixel brightness of the images. The values, as expected of a trained Model, not only lie in same range, but lie close to the diagonal and show an approximate equality. The centroids of both the ground truth images and the predicted images lie within small ranges of pixel values as should be expected of an ideally trained model. Please see the relevant text for detailed comments.

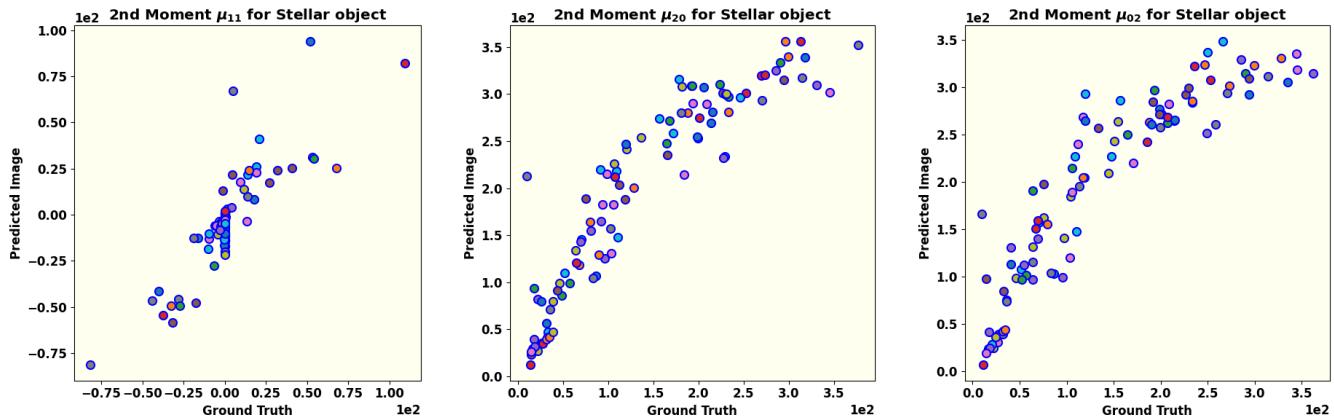


Figure 11. The second order central moment of the brightness distribution of the images is analogous to the moment of inertia of a mass distribution. The panels of this figure show the comparison of the second order moments ( $\mu_{11}$ ,  $\mu_{20}$ , and  $\mu_{02}$ ) of the predicted images vs. the 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.

An ideally trained Model should predict the overall pixel brightness of the generated images to be in the same range of values, if not equal to those of the ground truth images. The left panel of Fig. 10 displays the monopole values for 100 reconstructed images. The plot reveals a broad agreement in the overall intensity of the ground truth targets and the predicted (generated) images. However, a slightly higher brightness of reconstructed images in the intermediate range of brightness is evident. This could, perhaps, be mitigated by tuning the Batch Size hyperparameter.

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

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

$$y_c = \frac{M_{01}}{M_{00}} = \frac{\sum_{x,y} y \cdot I(x, y)}{\sum_{x,y} I(x, y)} \quad (19)$$

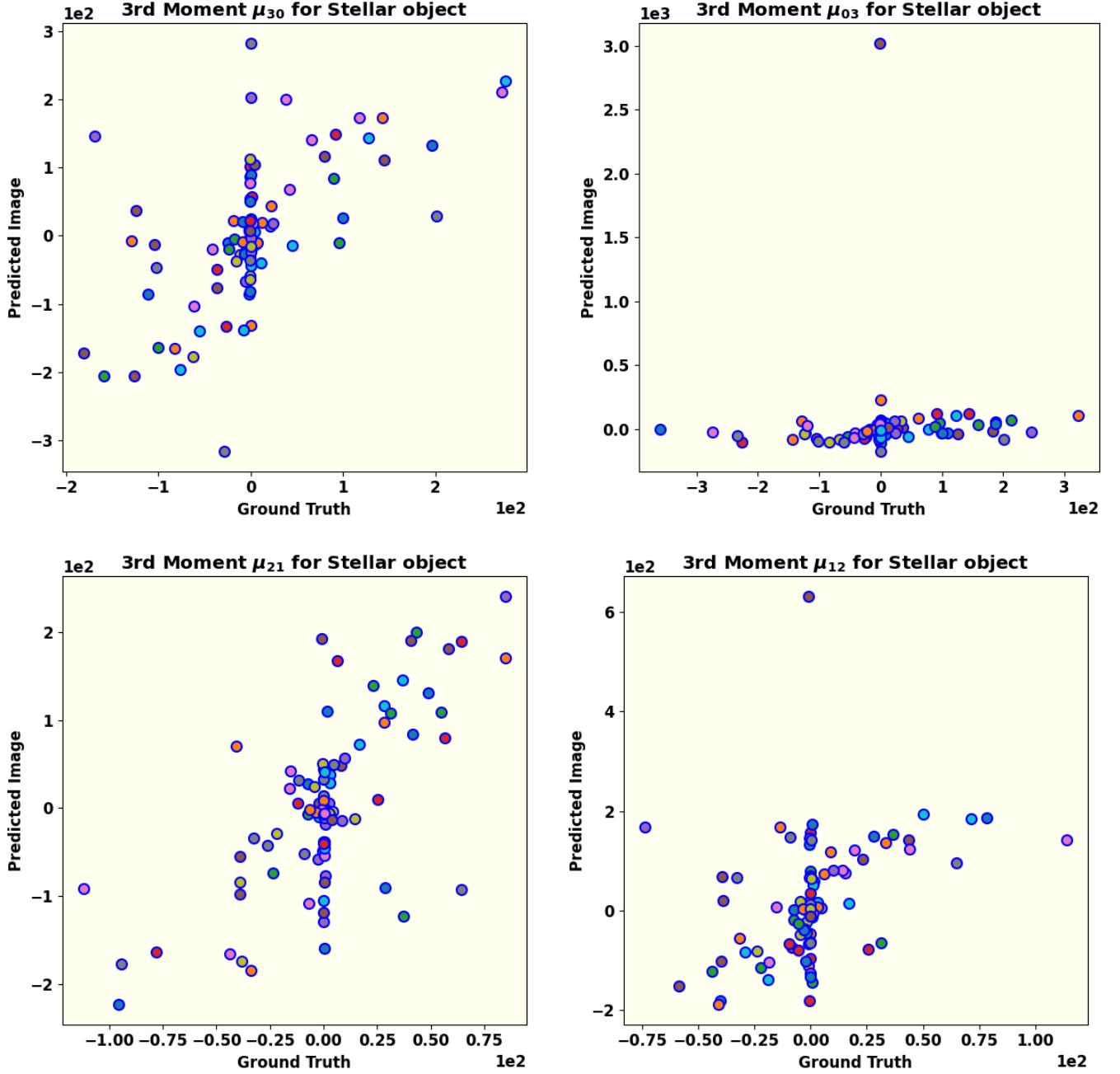


Figure 12. Shown here are all the third-order central moments ( $\mu_{30}$ ,  $\mu_{03}$ ,  $\mu_{21}$ , and  $\mu_{12}$ ) for ground truth and predicted images generated by the trained Model. They represent the skewness of the brightness distributions. The low skewness in the brightness distribution of the ground truth images is evident. One outlier among the predicted images dominates their skewness distribution along the  $y$ -axis as seen in the values of  $\mu_{03}$  and  $\mu_{12}$ . The trained Model, however, seems to have learned the approximately uniform distribution of skewness along the  $x$ -axis as is evident, with a few outliers though, in the values of  $\mu_{30}$  and  $\mu_{21}$ .

represent the centroids  $(x_c, y_c)$  of the pixel brightness distribution of the images. An ideally trained GAN model should produce the centroids of the target and predicted images within a small range of pixels. A larger range would imply a low fidelity in learning the brightness distribution of ground truth images. The middle and right panels of Fig. 10 compare the centroids  $(x_c, y_c)$  of 100 predicted images with their corresponding ground truths respectively. We notice that the coordinates of the centroids of ground truth images are spread across a range of 2 pixels whereas those of the predicted images are spread across a span of 4 pixels along  $x$ -axis and 7 pixels along  $y$ -axis. This clustering of centroids within a small scale range across all results indicates that the reconstructed images never have large offsets.

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. 11 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 third-order central moments**  $(\mu_{30}, \mu_{03}, \mu_{21}, \mu_{12})$  of the images quantify the degree and direction of asymmetry of the brightness distribution around their respective centroids. Fig. 12 presents a comparison of all third-order moments for both the ground truth and the reconstructed image. All the four panels show a low skewness in the brightness distribution of the ground truth images. The skewness of the predicted images along the  $y$ -axis is dominated by one outlier as seen in the values of  $\mu_{03}$  and  $\mu_{12}$ . However, the approximately uniform

distribution of skewness of the predicted images along the  $x$ -axis, with a few outliers though, as seen in the values of  $\mu_{30}$  and  $\mu_{21}$  suggests that this feature in the ground truth images has been picked up fairly well during the training of the Model.

### 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.

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

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

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

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

where,

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

and

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

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

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

Eqs. 21-25 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 the size, shape and brightness distribution of fast rotating stars can be recovered by a Pix2Pix cGAN model 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”. The Discriminator of our Model is trained to efficiently distinguish between the real images and fake (generated) images based on the “ground truth” images and the respective II data as the condition. The Generator is trained to generate progressively realistic images which are also consistent with the condition of the II data. The evaluation of the trained Model is then carried out by comparison of image moments of ground truth images and generated images. Specifically, the monopole, second, and third-order moments are compared. The results support the effectiveness of cGAN in achieving the stated objective.

While the results of this study highlight the significant potential of machine learning, and in particular the applicability of cGAN, for image reconstruction in II, several aspects require further refinement. First, an important factor in the reconstruction process is the extent of Fourier plane coverage, which depends on the number of available telescopes and the total observing time. The reasonable success of this piece of work suggests that a network of higher number of telescopes providing higher

number of baselines and greater coverage of the  $(u, v)$  plane signal, would play a critical role in projects of image reconstruction of more complicated stellar systems can be undertaken. Future work might explore different observatory layouts to assess their impact on image reconstruction quality. Second, detector efficiencies, which impact the signal-to-noise ratio (SNR) of actual observational data, have not yet been incorporated; addressing these factors will be crucial for more accurate SNR estimation. Third, exploring and comparing alternative methods for image generation could reveal approaches that outperform cGAN in reconstructing stellar images with II. Fourth, experimenting with different loss functions could provide additional insights into the reconstruction quality. Although further testing is needed to refine the GAN and enhance its robustness and reliability, our findings suggest that machine learning is a promising approach for phase reconstruction in II.

#### NOTE ON SOFTWARE

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

#### ACKNOWLEDGEMENTS

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

#### REFERENCES

- 1157 Abadi et al., M. 2016, arXiv preprint arXiv:1603.04467
- 1158 Abe et al., S. 2024, MNRAS, 529, 4387
- 1159 Acciari et al., V. A. 2020, MNRAS, 491, 1540
- 1160 Acharyya et al., A. 2024, The Astrophysical Journal, 966, 28
- 1162 Aleksić et al., J. 2016, Astroparticle Physics, 72, 76
- 1163 Archer, A., Aufdenberg, J. P., Bangale, P., et al. 2025, arXiv e-prints, arXiv:2506.15027, doi: [10.48550/arXiv.2506.15027](https://doi.org/10.48550/arXiv.2506.15027)
- 1166 Baumgartner et al., S. 2020, MNRAS, 498, 4577
- 1167 Coccolini et al., D. 2021, arXiv preprint arXiv:2122.11578
- 1169 Domiciano de Souza, A., Kervella, P., Jankov, S., et al. 2003, Astronomy and Astrophysics, 407, L47, doi: [10.1051/0004-6361:20030868](https://doi.org/10.1051/0004-6361:20030868)
- 1171 Domiciano de Souza, A., Kervella, P., Jankov, S., et al. 2005, Astronomy and Astrophysics, 442, 567, doi: [10.1051/0004-6361:20042476](https://doi.org/10.1051/0004-6361:20042476)
- 1174 Dravins et al., D. 2013, Astroparticle Physics, 43, 331, doi: [10.1016/j.astropartphys.2012.04.017](https://doi.org/10.1016/j.astropartphys.2012.04.017)
- 1176 Fienup, J. 1982, Applied Optics, 21, 2758
- 1177 Gamo, H. 1963, Journal of Applied Physics, 34, 875
- 1178 Gerchberg, R. W. 1972, Optik, 35, 237
- 1179 Glauber, R. J. 1963, Physical Review, 130, 2529
- 1180 Goldberger et al., M. L. 1963, Physical Review, 132, 2764
- 1182 Goodfellow et al., I. 2014, Advances in neural information processing systems, 27
- 1183 Hanbury Brown, R., & Twiss, R. Q. 1956, Nature, 177, 27
- 1184 Hanbury Brown et al., R. 1957, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 242, 300

- 1187 Hanbury Brown et al., R. 1958, Proceedings of the Royal  
 1188 Society of London. Series A. Mathematical and Physical  
 1189 Sciences, 243, 291
- 1190 Hanbury Brown et al., R. 1974, Monthly Notices of the  
 1191 Royal Astronomical Society, 167, 121
- 1192 Haubois et al., X. 2009, Astronomy & Astrophysics, 508,  
 1193 923, doi: [10.1051/0004-6361/200912927](https://doi.org/10.1051/0004-6361/200912927)
- 1194 Holmes et al., R. 2010, in Adaptive Coded Aperture  
 1195 Imaging, Non-Imaging, and Unconventional Imaging  
 1196 Sensor Systems II, Vol. 7818, SPIE, 175–185
- 1197 Hu, M.-K. 1962, IRE transactions on information theory, 8,  
 1198 179
- 1199 Hunter, J. D. 2007, Computing in Science & Engineering, 9,  
 1200 90, doi: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55)
- 1201 Isola et al., P. 2017, in Proceedings of the IEEE conference  
 1202 on computer vision and pattern recognition, 1125–1134
- 1203 Kirisits et al., C. 2024, arXiv preprint arXiv:2406.14143v2
- 1204 Le Bohec, S; Holder, J. 2006, The Astrophysical Journal,  
 1205 649, 399
- 1206 Li et al., X. 2014, in Imaging Spectroscopy, Telescopes and  
 1207 Large Optics, Vol. 9298, SPIE, 92981G1–G7
- 1208 Liu, L.-C., Wu, C., Li, W., et al. 2024, arXiv preprint,  
 1209 arXiv:2404.15685, doi: [10.48550/arXiv.2404.15685](https://doi.org/10.48550/arXiv.2404.15685)
- 1210 Liu et al., L.-C. 2025, Physical Review Letters, 134, 180201,  
 1211 doi: [10.1103/PhysRevLett.134.180201](https://doi.org/10.1103/PhysRevLett.134.180201)
- 1212 Lucy, L. B. 1967, Zeitschrift für Astrophysik, Vol. 65, p. 89,  
 1213 65, 89
- 1214 Maeder, A. 1999, Astronomy & Astrophysics, 347, 185
- 1215 Martinez, A. O., Baron, F., Monnier, J. D., Roettenbacher,  
 1216 R. M., & Parks, J. R. 2021, The Astrophysical Journal,  
 1217 916, 60, doi: [10.3847/1538-4357/ac06a5](https://doi.org/10.3847/1538-4357/ac06a5)
- 1218 McAlister et al., H. A. 2005, The Astrophysical Journal,  
 1219 628, 439
- 1220 Mirza et al., M. 2014, arXiv preprint arXiv:1411.1784
- 1221 Monnier, J. D., Zhao, M., Pedretti, E., et al. 2007, Science,  
 1222 317, 342, doi: [10.1126/science.1143205](https://doi.org/10.1126/science.1143205)
- 1223 Murphy, K. P. 2022, Probabilistic machine learning: an  
 1224 introduction (MIT press)
- 1225 Mustafa et al., M. 2019, Computational Astrophysics and  
 1226 Cosmology, 6, 1
- 1227 Norris et al., R. P. 2021, The Astrophysical Journal, 919,  
 1228 124, doi: [10.3847/1538-4357/ac0c7e](https://doi.org/10.3847/1538-4357/ac0c7e)
- 1229 Nuñez et al., P. D. 2010, in Optical and Infrared  
 1230 Interferometry II, Vol. 7734, SPIE, 458–467
- 1231 Nuñez et al., P. D. 2012a, Monthly Notices of the Royal  
 1232 Astronomical Society, 419, 172
- 1233 Nuñez et al., P. D. 2012b, Monthly Notices of the Royal  
 1234 Astronomical Society, 424, 1006
- 1235 Pedretti, E., Monnier, J. D., ten Brummelaar, T., &  
 1236 Thureau, N. D. 2009, New Astronomy Reviews, 53, 353,  
 1237 doi: [10.1016/j.newar.2010.07.005](https://doi.org/10.1016/j.newar.2010.07.005)
- 1238 Prince, S. J. 2023, Understanding deep learning (MIT  
 1239 press)
- 1240 Rai, K. N., Basak, S., Sarangi, S., & Saha, P. 2025,  
 1241 Resonance, 30, 45, doi: [10.1007/s12045-025-1729-x](https://doi.org/10.1007/s12045-025-1729-x)
- 1242 Rai et al., K. N. 2021, Monthly Notices of the Royal  
 1243 Astronomical Society, 507, 2813,  
 1244 doi: [10.1093/mnras/stab2391](https://doi.org/10.1093/mnras/stab2391)
- 1245 Rai et al., K. N. 2022, Monthly Notices of the Royal  
 1246 Astronomical Society, 516, 2864,  
 1247 doi: [10.1093/mnras/stac2433](https://doi.org/10.1093/mnras/stac2433)
- 1248 Ronneberger et al., O. 2015, in Medical image computing  
 1249 and computer-assisted intervention—MICCAI 2015: 18th  
 1250 international conference, Munich, Germany, October 5–9,  
 1251 2015, proceedings, part III 18, Springer, 234–241
- 1252 Sato et al., T. 1978, Applied optics, 17, 2047
- 1253 Sato et al., T. 1979, Applied Optics, 18, 485
- 1254 Sato et al., T. 1981, Applied Optics, 20, 2055
- 1255 Schawinski et al., K. 2017, MNRAS, 467, L110
- 1256 Teague, M. R. 1980, Josa, 70, 920
- 1257 Teague, M. R. 1983, Journal of Opticala Society of  
 1258 America, 73, 1434
- 1259 van Belle, G. T., Ciardi, D. R., Thompson, R. R., Akeson,  
 1260 R. L., & Lada, E. A. 2001, The Astrophysical Journal,  
 1261 559, 1155, doi: [10.1086/322340](https://doi.org/10.1086/322340)
- 1262 Virtanen et al., P. 2020, Nature methods, 17, 261
- 1263 Vogel et al., N. 2025, MNRAS, 537, 2334
- 1264 Von Zeipel, H. 1924, Monthly Notices of the Royal  
 1265 Astronomical Society, Vol. 84, p. 665–683, 84, 665
- 1266 Zhang et al., J. 2020, Opt. Lett., 45, 3649
- 1267 Zhao, M., Monnier, J. D., Pedretti, E., et al. 2009, The  
 1268 Astrophysical Journal, 701, 209,  
 1269 doi: [10.1088/0004-637X/701/1/209](https://doi.org/10.1088/0004-637X/701/1/209)