



University of
Zurich^{UZH}

Research internship astro (particle) physics

*Approaching phase retrieval in
Intensity Interferometry with
Imaging Atmospheric Cherenkov
Telescopes using Generative
Adversarial Networks
(provisional)*

by
Yuri van der Burg
at the
Department of Physics
University of Zurich

Supervisor
Prof. Prasenjit Saha

October 22, 2023

Abstract

In this research project, we apply machine learning methods to the phase retrieval problem that arises in Intensity Interferometry. A conditional GAN is used to reconstruct the shape, size, and brightness distribution of a fast-rotating star, based on its sparsely sampled phase-subtracted Fourier transform. This project focuses on Intensity Interferometry with IACTs located in La Palma, Spain, but the principles can be applied to any setup. Assuming a full night of observation, a basic reconstruction can already be achieved using the two MAGIC telescopes. Adding two LSTs increases the reconstruction of the brightness distribution greatly. With seven sub-mirrors on one LST, optimal reconstruction can be achieved with 6 hours of observation.

This work acts as a proof of concept that machine learning can be reliably used in Intensity Interferometry, opening the possibility of obtaining images of yet unresolved astronomical objects.

Contents

1	Introduction	3
2	Theory	4
2.1	Intensity Interferometry	4
2.2	Intensity Interferometry with MAGIC	4
2.3	Baseline considerations	6
2.4	Generative Adversarial Networks	7
2.4.1	Generator	8
2.4.2	Discriminator	8
3	Setup	9
3.1	Data preparation	9
3.2	GAN architecture	11
3.3	Hyperparameter tuning	13
3.4	Robustness	18
3.5	LST sub-mirrors	21
4	Results	22
4.1	Phase retrieval	22
4.2	Robustness	26
4.3	LST sub-mirrors	26
5	Conclusion and outlook	31

1 Introduction

In the years after Intensity Interferometry was first carried out for stellar objects in the 1970s, it has not received a lot of attention. Increased time resolution of the electronics allows Intensity Interferometry to be a viable option to resolve astrophysical objects at visible wavelengths. Currently, Intensity Interferometry with IACTs is used to measure the angular diameter of stars [1]. With the construction of CTA, the number of telescopes that could be available for Intensity Interferometry increases, which allows us to probe a larger area of the signal in the Fourier plane. A full reconstruction of the brightness distribution of the source requires phase reconstruction, as Intensity Interferometry measures the complex visibility squared $|V_{12}|^2$, in contrast to Amplitude Interferometry, which measures V_{12} . This work approaches the phase reconstruction using a conditional Generative Adversarial Network on simulated images of fast-rotating stars.

The principles of Intensity Interferometry and Generative Adversarial Networks are explained in chapter 2. Chapter 3 explains how the data was simulated, processed, and used in the network. It also includes the specific Generative Adversarial Network architecture, as well as the hyperparameter tuning. The results are reported in chapter 4, followed by a conclusion and outlook.

2 Theory

This chapter starts with a short overview of Intensity Interferometry, followed by an explanation of how Intensity Interferometry is done with Cherenkov Imaging Telescopes. In a second part follows a description of Generative Adversarial Networks.

2.1 Intensity Interferometry

Intensity Interferometry (II) was first carried out by R. Hanbury Brown with the Narrabri Stellar Intensity Interferometer, the first optical experiment that resolved the diameter of bright stars at visible wavelengths [2, 3].

In a simple case, an intensity interferometer consists of two light collectors. This can be two optical telescopes, for example, MAGIC-I and MAGIC-II, which are pointed to a star and measure each an intensity $I_1(t), I_2(t)$ [4, 5]. The signals from both detectors are cross-correlated and measured over time. The second-order coherence function is given by:

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

Where τ is the time delay between both telescopes. If the light source is chaotic and randomly polarized, we have:

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

Here, Δf is the electronic bandwidth, $\Delta \nu$ is the optical bandwidth and we have $\Delta f \ll \Delta \nu$. V_{12} is the complex visibility, the Fourier transform of the source brightness distribution [4]. The phase information of V_{12} can not be measured, hence it is lost.

The correlation is often expressed in terms of the normalized contrast, given by:

$$c = \frac{\langle (I_1(t) - \langle I_1 \rangle) \cdot (I_2(t + \tau) - \langle I_2 \rangle) \rangle}{\langle I_1(t) \rangle \cdot \langle I_2(t) \rangle} = g^{(2)} - 1 \quad (3)$$

2.2 Intensity Interferometry with MAGIC

The two MAGIC (Major Atmospheric Gamma Imaging Cherenkov) telescopes are two Imaging Atmospheric Cherenkov Telescopes (IACTs) located at the Roque de los Muchachos Observatory in La Palma, Spain. The main physics channel of MAGIC is the study of Very High Energy gamma rays

(≥ 30 GeV) originating from particle showers in the atmosphere. Both telescopes have an array of mirrors with a total diameter of 17m, which focuses the light onto an array of photomultiplier tubes (PMTs) [6]. For Intensity Interferometry, the light of a stellar source is focused on a single PMT, which has a filter in front. The purpose of this filter is to efficiently transmit light around 432nm, which protects the PMTs from excessive light. This is necessary, as II observations are mainly done during full moon nights, where it is too bright for Very High Energy γ -ray observations. The optical signal of the PMTs is transmitted with optical fibers, converted to an electrical signal and digitized [4].

The measurable observable is the Pearson's correlation coefficient:

$$\rho(\tau) = \frac{\langle (I_1(t) - \langle I_1 \rangle) \cdot (I_2(t + \tau) - \langle I_2 \rangle) \rangle}{\sqrt{\langle (I_1(t) - \langle I_1 \rangle)^2 \rangle} \cdot \sqrt{\langle (I_2(t) - \langle I_2 \rangle)^2 \rangle}} \quad (4)$$

Since I_1 and I_2 are proportional to the direct current of the PMTs (DC_i), the normalized contrast can be calculated as in equation 5. One must also correct for a non-constant gain of the PMTs, as well as the moon. The latter can be done by measuring the background light with an additional PMT with no mirrors focused on it [4].

$$c \propto \frac{\rho}{\sqrt{(DC_1 \cdot DC_2)}} \quad (5)$$

It must be noted that the calculation of Pearson's correlation coefficient is non-trivial: MAGIC makes use of the Convolution Theorem for discrete Fourier transforms, because it is computationally more efficient.

The significance of the signal can be expressed as the signal-to-noise ratio S/N , which depends on a lot of factors, but most importantly it is proportional to the complex visibility, which itself depends on the distance between the telescopes, shown in equation (6) [4].

$$S/N = A \cdot \alpha \cdot q \cdot n \cdot |V_{12}|^2(d) \cdot \sqrt{b_v} \cdot F^{-1} \cdot \sqrt{\frac{T}{2}} \cdot \sigma \quad (6)$$

Here, A is the total mirror area, α the quantum efficiency of the PMTs, q the quantum efficiency of the optics, n the differential photon flux from the source, and b_v the cross-correlation bandwidth. The noise of the PMTs is accounted for with F , T denotes the observation time, and σ is the normalized spectral distribution of the light (including filters) [4]. While most of the parameters can be optimized with hardware, the only way to obtain a better signal-to-noise ratio is to increase the observation time.

2.3 Baseline considerations

Since the complex visibility depends on the distance between the telescopes, also called the baseline d , one can measure the angular size of the star using equation (7). The idea is to measure with different baselines, such that the size of the star can be fitted. This has been done for several examples [4, 7].

$$V_{12}(d) = \frac{c(d)}{c(0)} \quad (7)$$

If the source is at the zenith, that is perpendicular to the line of sight, the coordinates in the Fourier plane (u, v) are given by:

$$(u, v) = \frac{1}{\lambda}(d_N, d_E) \quad (8)$$

Where B_N and B_E are the baseline coordinates expressed in north and east coordinates. Since not all sources are at the zenith, and the telescopes are stationary, the source moves across the sky during the observation. The baseline rotation is then given by equation (9), which traces an ellipse for every pair of telescopes. Furthermore, the different altitudes d_A of the telescopes must be considered [5, 8].

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = R_x(\delta) \cdot R_y(h) \cdot R_x(-l) \begin{pmatrix} d_N \\ d_E \\ d_A \end{pmatrix} \quad (9)$$

Here, δ is the declination and h is the hour angle of the stellar source, and l is the latitude of the telescopes. The three matrices R_i correspond to the fundamental representation of the SO(3) group [8].

Since every pair of telescopes traces an ellipse in the Fourier plane, the total number of ellipses scales as:

$$\mathcal{N} = \frac{N_T \cdot (N_T - 1)}{2} \quad (10)$$

Because the number of baselines increases non-linearly, Intensity Interferometry benefits greatly from a large number of telescopes, for example as it is planned with the Cherenkov Telescope Array (CTA). At the time of this work, one of the four Large Sized Telescopes (LSTs) in the northern CTA is in operation, three additional LSTs are in construction.

2.4 Generative Adversarial Networks

Generative Adversarial Networks (GANs) were invented by Ian Goodfellow in 2014. The concept is rather simple, as it consists of two competing networks. The first network, called the generator, creates new images based on an input image. For consistency, the images generated by the generator will be denoted as generated images. The second network, the discriminator, tries to distinguish between the generated image and the real image, also called the ground truth [9]. As a consequence of training both networks alternately, the generated images eventually become indistinguishable from the real images. This is nothing than a two-player min-max game, a famous problem in game theory. Goodfellow initially proposed the problem as:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log (1 - D(G(z)))] \quad (11)$$

Here, $V(D, G)$ is the value function of the min-max game. The goal is now to learn the distribution of the generator p_g over the data x . We have input noise variables $p_z(z)$, as well as the two perceptrons $G(z; \theta_g)$ and $D(x; \theta_d)$ with parameters θ_i . G represents a differential function that maps from z to the data space x , while $D(x)$ represents the probability that x is from the real data [9].

The problem can be reformulated as:

$$\max_D V(G, D) = \mathbb{E}_{x \sim p_{data}} [\log D_G^*(x)] + \mathbb{E}_{x \sim p_g} [\log (1 - D_G^*(x))] \quad (12)$$

Where D_G^* denotes the optimum of the discriminator for a given fixed generator, shown in equation (13). It can now be shown that the global optimum of equation (12) is reached if and only if $p_g = p_{data}$. Furthermore, if both G and D are allowed to reach their optima, p_g converges to p_{data} . A more detailed discussion of the problem, including proofs can be found in [9].

$$D_G^*(x) = \frac{p_{data}(x)}{p_{data}(x) + p_g(g)} \quad (13)$$

GANs can be extended to a conditional model [10]. In this case, both generator and discriminator receive extra information y , such that the value function of the conditional GAN (cGAN) can be expressed as:

$$V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log D(x|y)] + \mathbb{E}_{z \sim p_z(z)} [\log (1 - D(G(z|y)))] \quad (14)$$

Furthermore, [11] observed that combining the cGAN given in equation 14 with the traditional L1-loss improves results, as the generator tries to be close to the ground truth. Hence the function, that is minimized is:

$$L_{tot} = \arg \min_G \max_D V(D, G) + \lambda \cdot L_1(G) \quad (15)$$

Where $L_1(g)$ is given by equation (16), and $\lambda = 100$.

$$\mathbb{E}_{x,y,z} [\|y - G(x, z)\|_1] \quad (16)$$

This type of network is shown to be very robust, as it has been applied to several problems. Examples include creating colored images from grayscale images, creating images of facades based on architectural labels, changing from day to night in different pictures, and predicting a map based on satellite data. A longer list of applications is given in [11].

2.4.1 Generator

The generator is a common U-Net convolutional network, described in [12]. It consists of a contracting part, reducing the image size, and an expansive part, where the image is enlarged. The contracting part consists of repeated application of alternating convolutional layers and rectified linear unit (ReLU) layers. The expansion consists of repeated convolution, batch normalization, and ReLU layers. The first contracting step also includes a dropout layer [11]

2.4.2 Discriminator

The discriminator, named PatchGAN, does not work globally but rather classifies single patches as real or fake. The overall result is then estimated based on the average. The model architecture is as follows: zero padding, convolution, batch-norm, ReLU, zero-padding, and convolution [11].

3 Setup

This section explains the workflow of this work, which consists of data preparation, GAN selection, and hyperparameter tuning. The code used for this work is made public and can be found on GitHub¹.

3.1 Data preparation

In the first step, images of fast-rotating stars are simulated in Python. The fast rotators are modeled as oblate spheroids with different radii and oblateness between 0.5 and 1. They are viewed from different angles. The effect of gravity darkening is modeled as a linear dependence, i.e. the brightness of the star increases linearly along the symmetry axis.

Figures 1 to 5 illustrate the corresponding steps, simulating a fast rotating star at the position of Altair with an oblateness of 0.7. Figure 1 shows the simulated star.

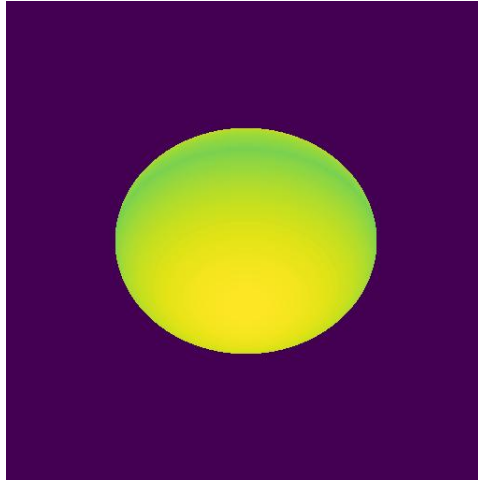


Figure 1: Simulated fast rotating star with an oblateness of 0.7. The brightness is highest at the pole in the front and there is gravitational darkening at the equator.

Because it is not possible to measure the full Fourier plane with Intensity Interferometry, and because the source in the sky moves concerning the telescopes, equation (9) must be applied. The traced ellipses arise when integrating over the hour angle of the source. For the hyperparameter tuning

¹https://github.com/Yurivanderburg/II_Project

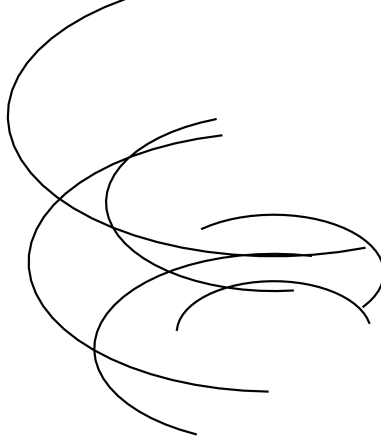
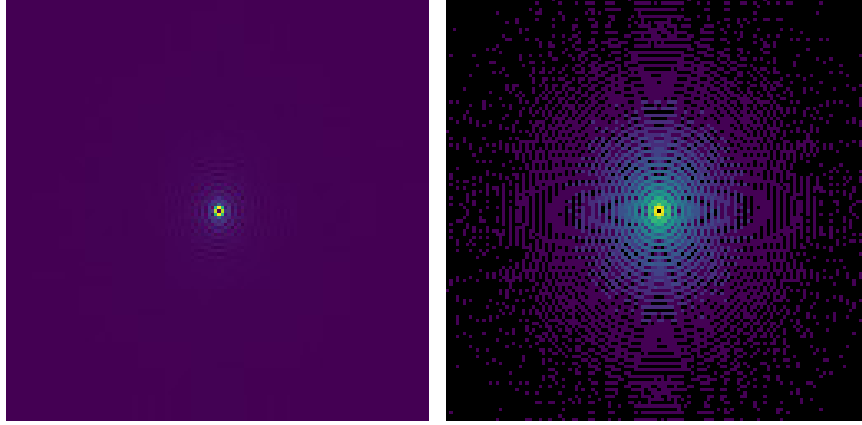


Figure 2: The six ellipses are traced by two MAGIC telescopes and two LSTs observing Altair for 11.5 hours. This is the sparse sampling map, which is the part of the Fourier plane that the model has access to.

and comparison of different telescopes, the total hour angle is about 11.5 hours. The ellipses are plotted and converted into grayscale images, resized, and stored as pure numpy arrays to simplify further analysis. The result is a map in the Fourier plane with several ellipses, also called the sparse sampling map. Figure 2 shows the sparse sampling for two MAGIC telescopes and two LSTs and Altair as the source.

To use the images in the GAN, the simulated rotating stars are further processed. First, Salt and Pepper noise is introduced at a rate of α (usually 0.5%). Then the images are resized and the mean is subtracted. A two-dimensional Fast Fourier Transform, as well as a Fourier shift, is applied, which yields a complex number of every pixel. Because Intensity Interferometry does not measure the phase, the absolute value is calculated. This is shown in figure 3a, denoted as Fourier plane throughout this work. To see the pattern of the Fourier transform, the same image in logarithmic scale is displayed in figure 3b.

This is also where the sparse sampling is introduced, namely by a pixel-wise multiplication between the absolute valued Fourier transformed image and the sparse sampling map. Finally, the pixels are normalized and converted to 8-bit integers. This image represents the sparsely sampled complex visibility $|V_{12}|$, as it can be measured with Intensity Interferometry. It is shown



(a) Linear scaling of the Fourier plane, as it is used in the model. (b) Logarithmic scaling of the Fourier plane, for display.

Figure 3: Absolute value of the two-dimensional Fast Fourier Transform of figure 1. This is what we measure in Intensity Interferometry, though we do not have access to the whole plane.

in figure 4, again with linear and logarithmic scaling.

The image shown in figure 4a is the input image for the GAN, which also requires the ground truth image. To reduce bias, the simulated stars are also resized with the same algorithm and converted to 8-bit integers. The GAN must know the ground truth corresponding to an input image, hence they are merged side-by-side, shown in figure 5. These images are what the GAN is trained with. This procedure is conducted for all the simulated stars, with 10% acting as test data, 10% as validation data, and the remaining 80% as training data.

3.2 GAN architecture

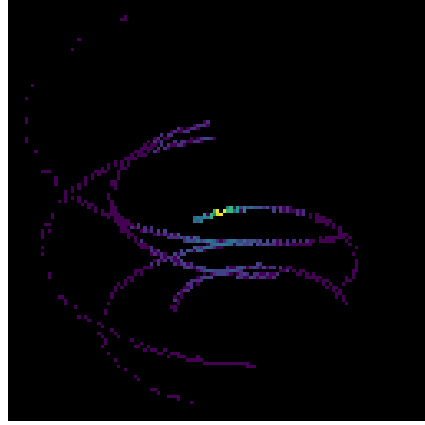
The GAN used in this work is based on pix2pix [11], which uses a cGAN as described in section 2.4. This architecture was selected, as it is very robust and could already be applied to various problems. An example can be found in the TensorFlow tutorials², where it is applied to a data set of architectural facades. To adapt the pix2pix GAN to the phase retrieval problem, some modifications are necessary.

The code used in this work is written in Python. The network uses the

²<https://www.tensorflow.org/tutorials/generative/pix2pix>



(a) Linear scaling of the sparsely sampled Fourier plane, as used in the GAN.



(b) Logarithmic scaling of the sparsely sampled Fourier plane, shown only for display.

Figure 4: Absolute valued FFT as in figure 3, probed by four telescopes. This is what we have access to, and what is used for the phase reconstruction.

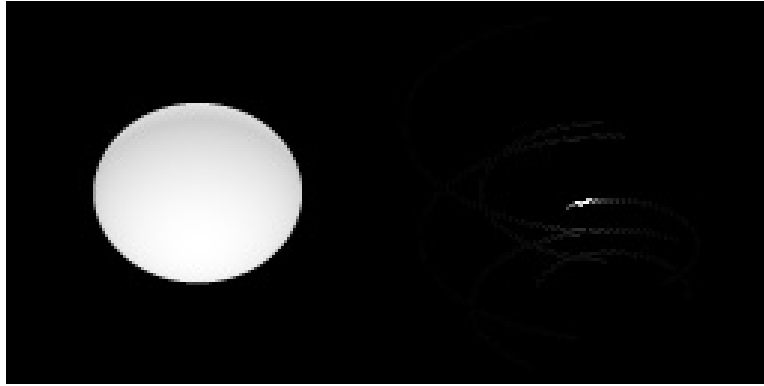


Figure 5: Merged image which includes the original and the sparsely sampled Fourier plane. This is exactly what the GAN receives.

TensorFlow [13] library, calculations are performed with scipy [14] and plots are drawn with matplotlib [15]. The models were trained on ScienceCluster, the computing facility of the University of Zurich.

3.3 Hyperparameter tuning

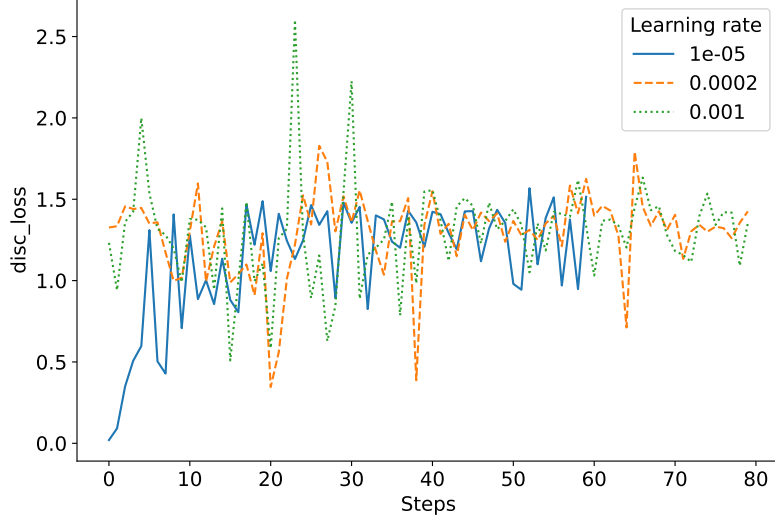
The Generative Adversarial Network used for this work depends on several parameters. The varied parameters will be explained here briefly. An in-depth discussion can be found in [16].

The learning rate of the optimizer is a measure of how much the model changes per iteration. Small learning rates can lead to underfitting, and large learning rates can make the model unstable. It is therefore important to carefully choose the learning rate [16]. Figure 6 shows the effect of the learning rate on both generator and discriminator loss for three different values of learning rate. Both loss functions have fewer outliers with decreasing learning rates, which is expected, as there is less change in the models with low learning rates. Even though all models stabilize at the same level, low learning rates are preferred.

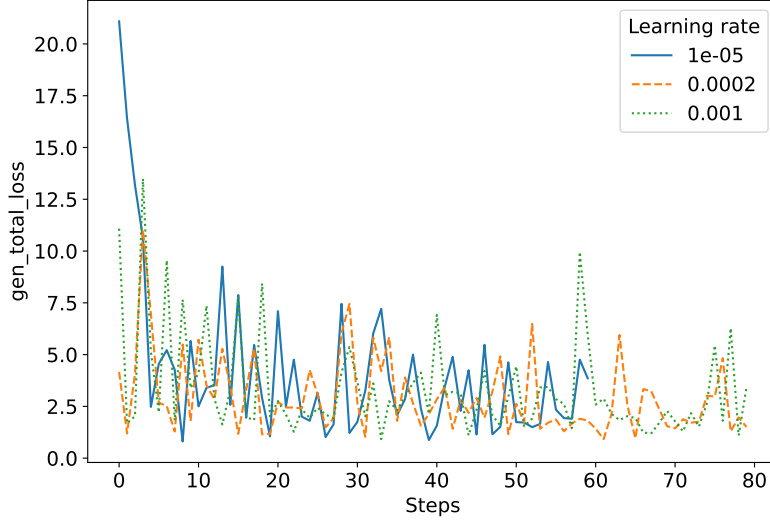
The kernel size is the size of the kernel in the convolutional layer. It determines the amount of pixels that are linearly combined into a new pixel. A large kernel size allows to catch features that are a few pixels away, but it also means that unrelated features might be introduced. The kernel size does not seem to have a big effect on the loss functions, as shown in figure 7. Smaller kernel sizes seem to have more outliers, which means that either generator or discriminator gains an advantage. Therefore, a kernel size of 5 is preferred.

The amount of noise is reflected by the two parameters α and β . Both indicate the percentage of pixels that change to either black or white, hence the name Salt and Pepper noise, where α is introduced in the real image, β in the generated image. Different ratios $\frac{\alpha}{\beta}$ can lead to different model performances. There is no significant difference between distinct noise rates introduced in the images. This is shown in figure 8, which shows the loss functions on smaller images (64 by 64). As there is no effect, it was not necessary to repeat this analysis for larger images.

The batch size defines the number of images that pass the network simultaneously. It is observed that smaller batch sizes can lead to better generalization [17]. Figure 9 shows two different batch sizes. When processing two images simultaneously, there seem to be fewer outliers, but since it increases the training time heavily, a batch size of 1 is used.

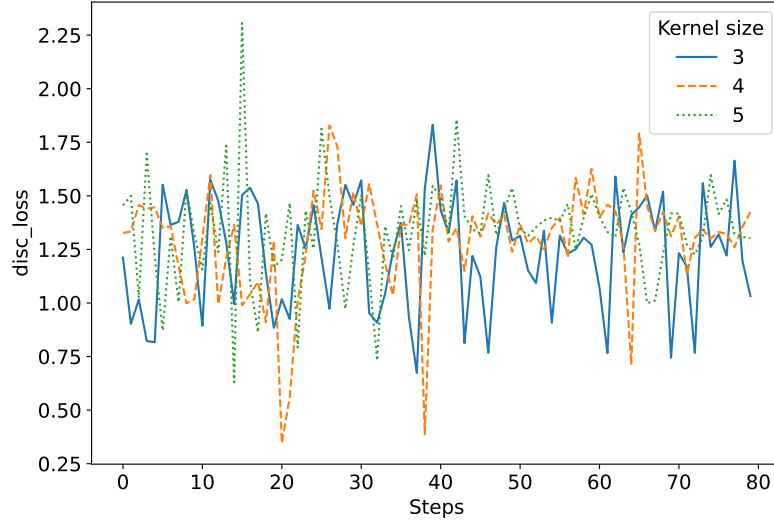


(a) Discriminator loss for three different learning rates.

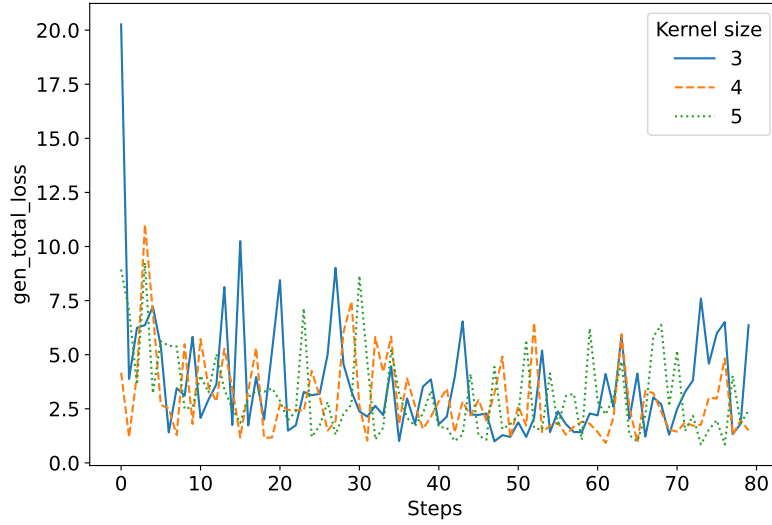


(b) Total generator loss for different learning rates, as in equation (15).

Figure 6: Generator and discriminator losses for three different learning rates. The upper figure shows the total discriminator loss, and the lower figure the total generator loss. There is no significant difference, but the highest learning rate might be prone to outliers.

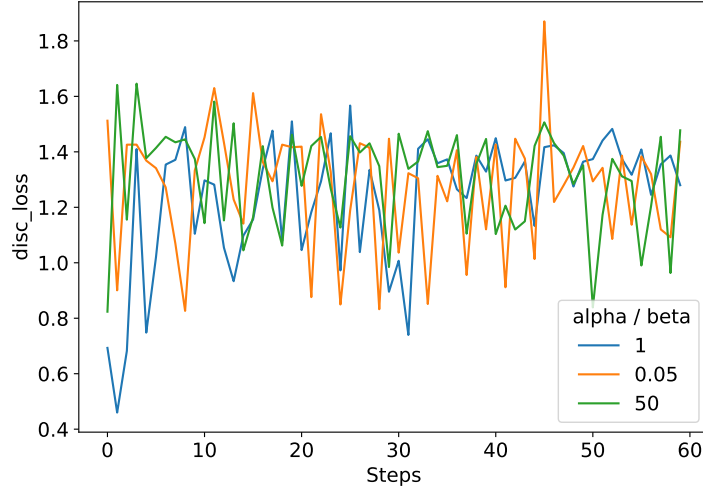


(a) Discriminator loss for three different kernel sizes.

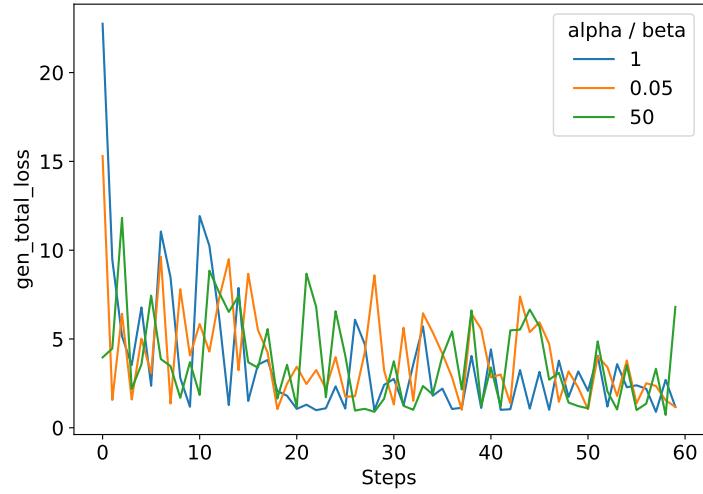


(b) Total generator loss for three different kernel sizes.

Figure 7: Generator and discriminator losses for three different kernel sizes in the convolutional layers. Here, the smallest kernel size has a lot of outliers, while the highest kernel size seems to be the most stable.

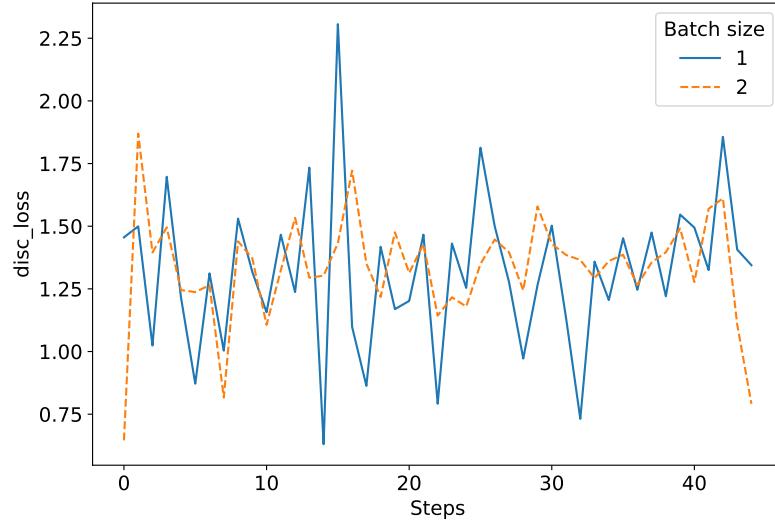


(a) Discriminator loss for different noise percentages.

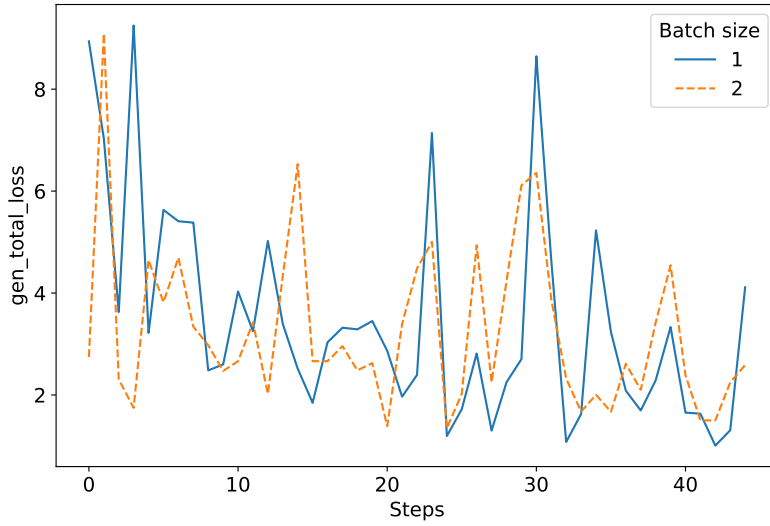


(b) Generator losses for different noise percentages.

Figure 8: Effect of the Salt and Pepper noise introduced in the images. Different ratios between α and β are shown. There is no significant effect. Please note that these results are from training on 64-pixel images.



(a) Discriminator loss for different batch sizes.



(b) Generator losses for different batch sizes.

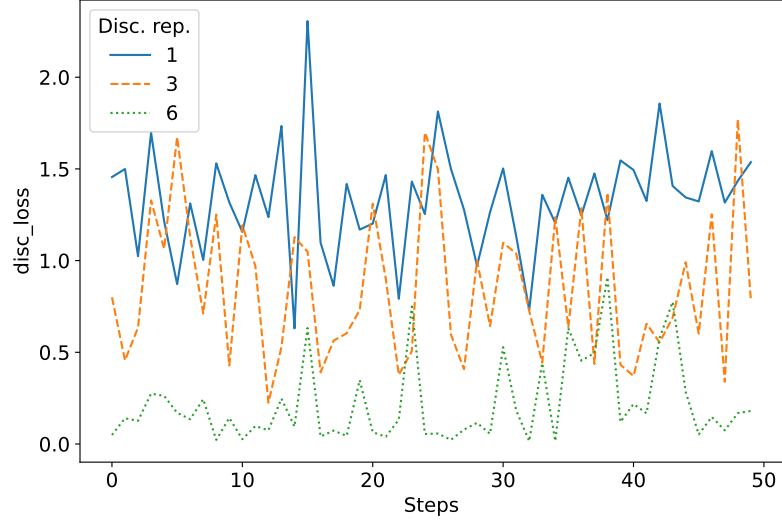
Figure 9: Loss functions for two different batch sizes. Large batch size seems to be more robust but also increases the training time heavily.

When training GANs, one can give the discriminator an advantage and increase the number of training steps of the discriminator before returning to the generator training. This can increase the performance of the model, but it increases the training time as well. It leads to a lower discriminator loss score, as shown in 10. At the same time, the generator score increases slightly, which should be no surprise. Because the generated images did not improve with higher discriminator training, it is usually trained the same amount as the generator.

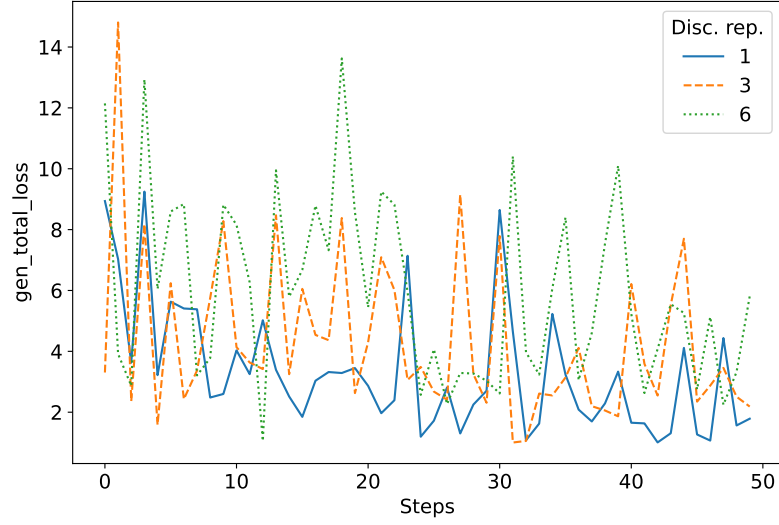
Finally, the amount of sparse sampling can be varied, giving the model access to more pixels. More telescopes automatically lead to more baselines and hence more pixels that are available. Figure 11 shows the loss functions for different numbers of telescopes. There is a huge disparity, which can partly be explained by the fact that the relation between telescopes and baselines is not linear. The Fourier plane can be sampled sixfold when using four telescopes in contrast to only two telescopes. In the latter case, both the generator and discriminator are not trained smoothly, as seen by the outliers. This is already better with three telescopes, and very promising when using four telescopes. The degree of sparse sampling seems to have the biggest effect of all hyperparameters.

3.4 Robustness

To estimate the robustness and the dependence on the sparse sampling map that is used in the training, one analysis includes six different sources. The procedure follows section 3.1, where five of those are selected to train the new model, leading to five times more training data. The last sparse sampling map is used only for testing, to estimate how well the model can adapt to a different sparse sampling.

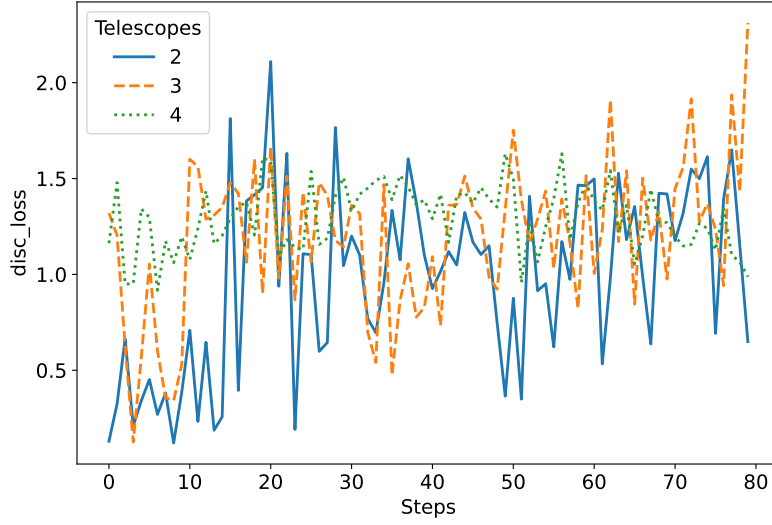


(a) Discriminator loss for different discriminator training.

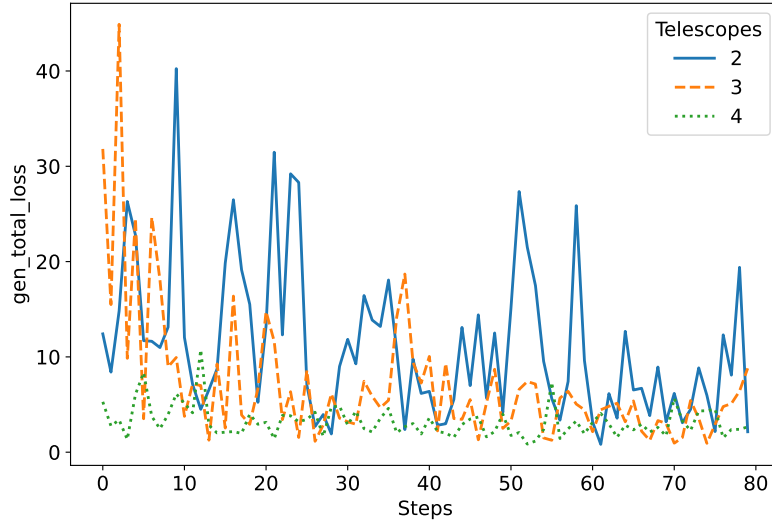


(b) Generator losses for different discriminator training.

Figure 10: The amount of discriminator training has a high impact on the discriminator score. The discriminator repetition indicates the factor by which the discriminator is trained more with respect to the generator.



(a) Discriminator loss for different amounts of telescopes.

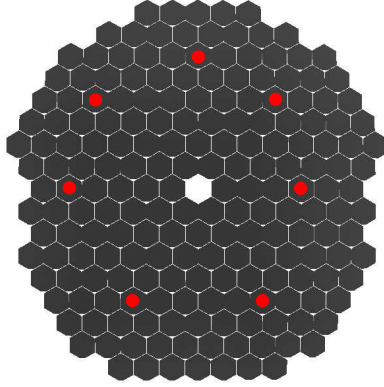


(b) Generator loss for different amount of telescopes.

Figure 11: The number of telescopes has a very impact on the model performance. If there are only two telescopes (1 baseline), both discriminator and generator are not trained smoothly: One gains a big advantage over the other. The result of four telescopes (6 baselines) is a lot better because the loss functions change only slightly with increasing steps.

3.5 LST sub-mirrors

Intensity Interferometry does not necessarily require two telescopes. Since we only need two photo sensors, it is possible to focus parts of the mirror on single PMTs. Again, the procedure follows 3.1, but instead of four telescopes, we use one LST with seven sub-mirrors, focused on Altair. The centers of the selected sub-mirrors are indicated in figure 12. The relative positions can be estimated with simple geometry, as the single facets are regular hexagons. The positions are given table 3.5, where d is the minimal diameter (side-to-side length) and D is the maximal diameter (long diagonal) of the hexagon. For LST, we have $d = 1.5\text{m}$ and $D = \frac{2}{\sqrt{3}} \cdot d \approx 1.73\text{m}$. This is done for three 3, 6, and 9 hours of observation.



Nr.	Position (x, y)
1	$(0, 4.5D)$
2	$(3d, 3D)$
3	$(4d, 0)$
4	$(2.5d, -3.75D)$
5	$(-2.5d, -3.75D)$
6	$(-5d, 0)$
7	$(-4d, 3D)$

Figure 12: Sketch of an LST mirror, where the centers of the sub-mirrors are indicated by red dots. Adapted from [18].

Table 1: Central position of the sub-mirrors on an LST. There is some asymmetry to avoid baseline redundancies.

4 Results

The best performance for image reconstruction has been observed with a learning rate of $2 \cdot 10^{-4}$, a kernel size of 5x5, and equal noise percentages in original and generated images. The batch size selected is 1, and discriminator and generator receive equal training. In this section, we will first discuss the phase retrieval using the above-mentioned hyperparameters, followed by an analysis where multiple sources are trained simultaneously. The chapter concludes with an analysis of LST sub-mirrors.

4.1 Phase retrieval

Figure 11b indicates a large difference in performance between different numbers of telescopes. The reason is intuitive: If the model has access to more pixels in the Fourier plane, more differences in pixel intensities can be used in the training, resulting in superior performance. The ground truth images (simulated fast-rotating star) and the generated images (reconstruction) for two, three, and four telescopes are displayed in figures 13, 14 and 15. More reconstructed images can be found on GitHub³.

As suspected, there is a clear difference in performance. If there is only one baseline (figure 13) the model is roughly able to estimate the size and the shape of the star, but the brightness distribution is not reliable.

Increasing the sparse sampling by a factor of three, we can already see an improvement, indicated in figure 14. Here, not only the size but also the shape of the star is reconstructed correctly. The brightness distribution, while not perfect, is good enough to estimate the orientation of the source. With four telescopes and six baselines, the size and shape of the source improve only a little, but the brightness distribution is much better, as shown in figure 15. This would allow us to study the star with high precision.

³https://github.com/Yurivanderburg/II_Project

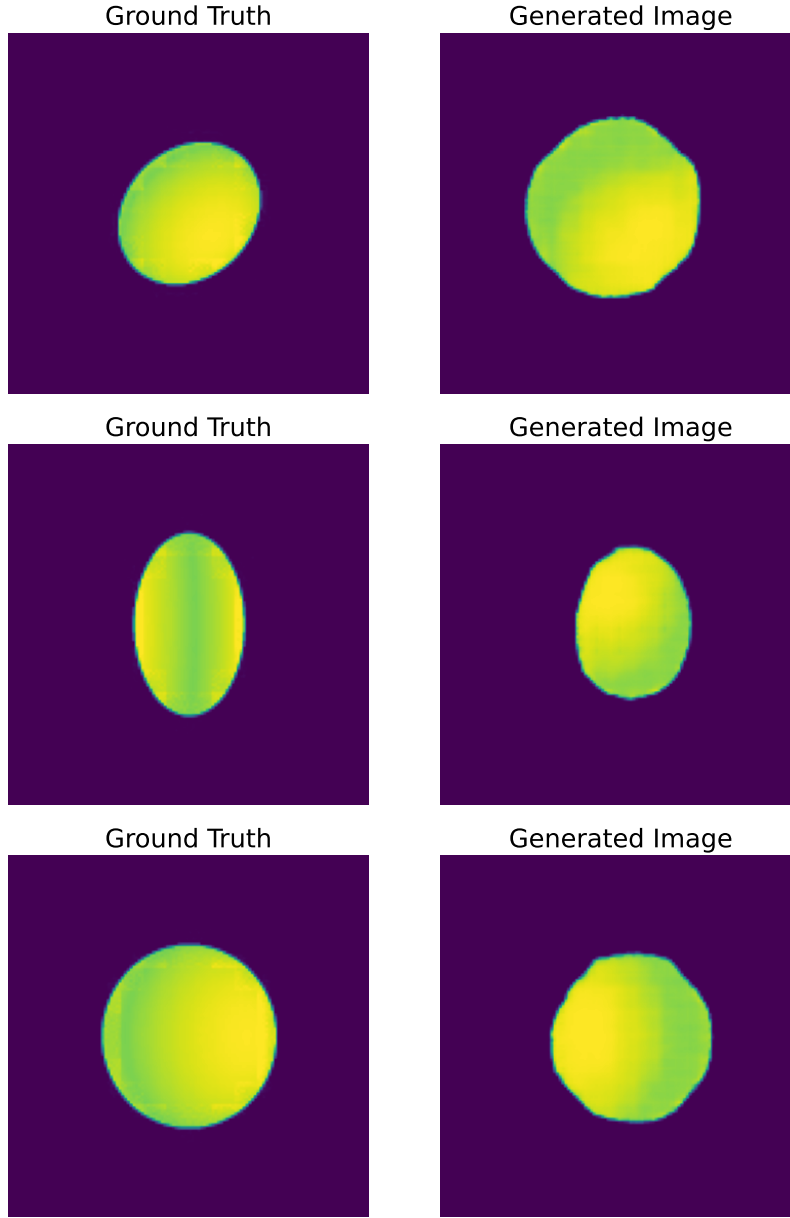


Figure 13: Simulated fast-rotating stars and the reconstructed images in a setup where two MAGIC telescopes observe a full night. While the reconstruction gives a rough estimation of the size and shape of the source, it fails to give detailed information.

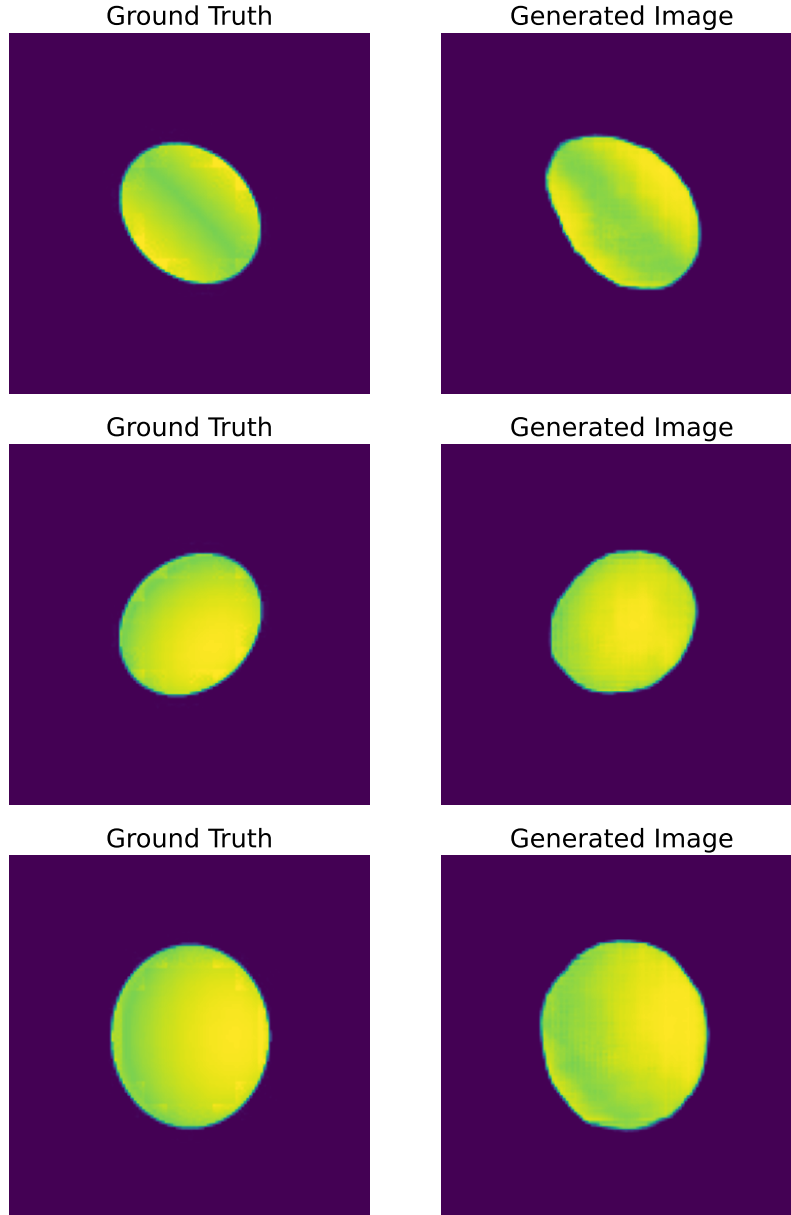


Figure 14: Simulated fast-rotating stars and the reconstructed images in a setup where two MAGIC telescopes and an LST observe a full night. Here, we can see that adding one telescope already improves the results greatly, in particular the brightness distribution is a lot more reliable.

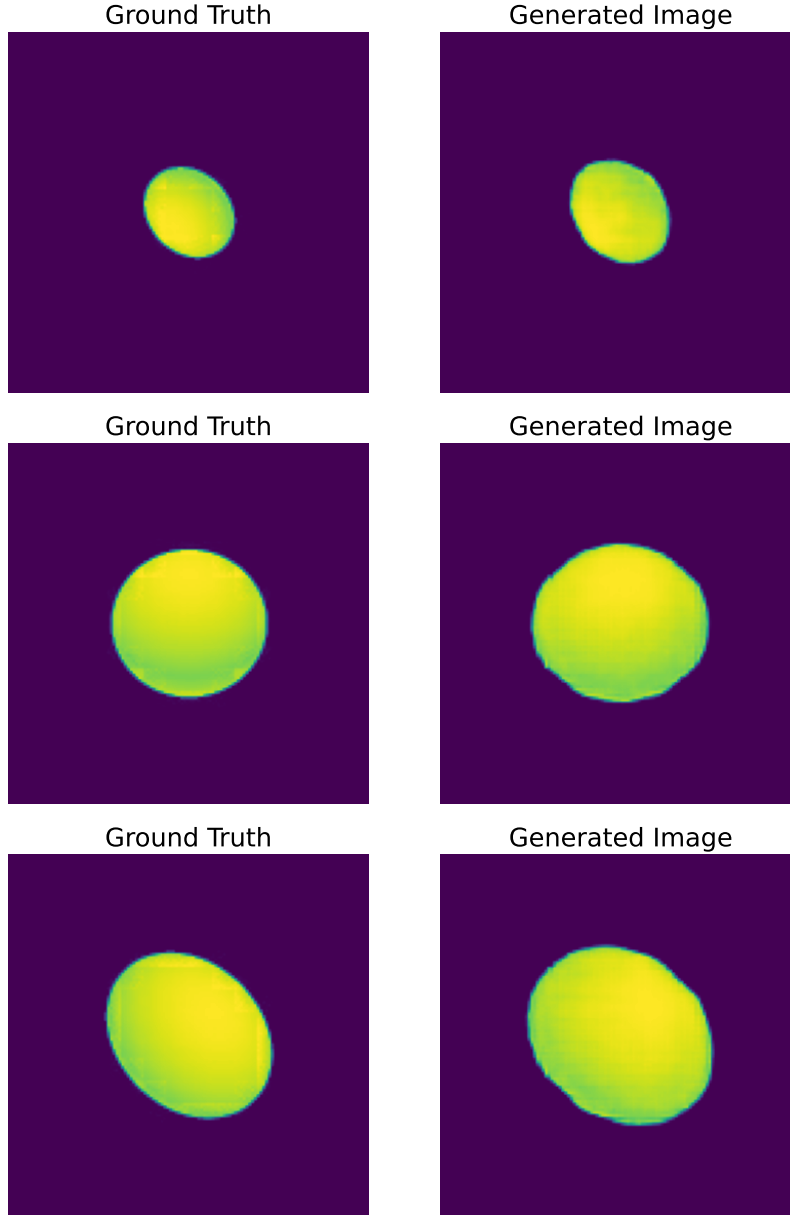


Figure 15: Simulated fast-rotating stars and the reconstructed images in a setup where two MAGIC telescopes and two LSTs observe a full night. The reconstruction is very advanced, as all features are correct.

4.2 Robustness

Figure 16 shows the image reconstruction of a model subject to unknown sparse sampling. We can see that in this case the reconstruction is not reliable, there is a bias toward inclined orientations. The model also fails to correctly calculate the size and shape of the star. This result is no surprise because the model needs to reconstruct the phase from an image for which it does not know how to interpret the pixel differences.

4.3 LST sub-mirrors

Figure 17 shows the reconstructed images for three hours of observation, figure 18 for six hours of observation, and 19 for nine hours. Again, we can see a difference in the quality of the reconstruction with increasing observation time. While three hours are enough to estimate the shape and size of the source, at least six hours are necessary to reconstruct the brightness distribution reliably. The difference between six and nine hours of observation is only minor compared to the difference between three and six hours, indicating that the model reaches an optimum.

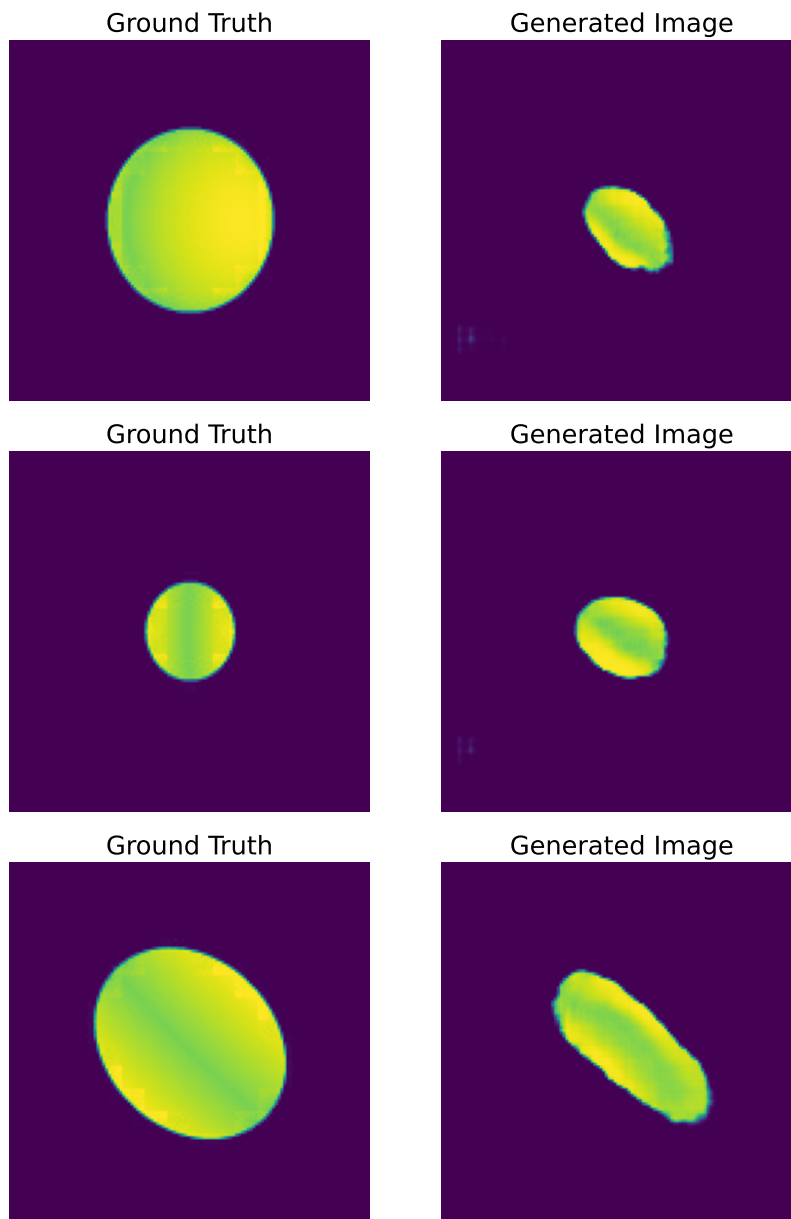


Figure 16: Image reconstruction of images whose sparse sampling map is not in the training dataset. We can see that this decreases the performance of the GAN massively.

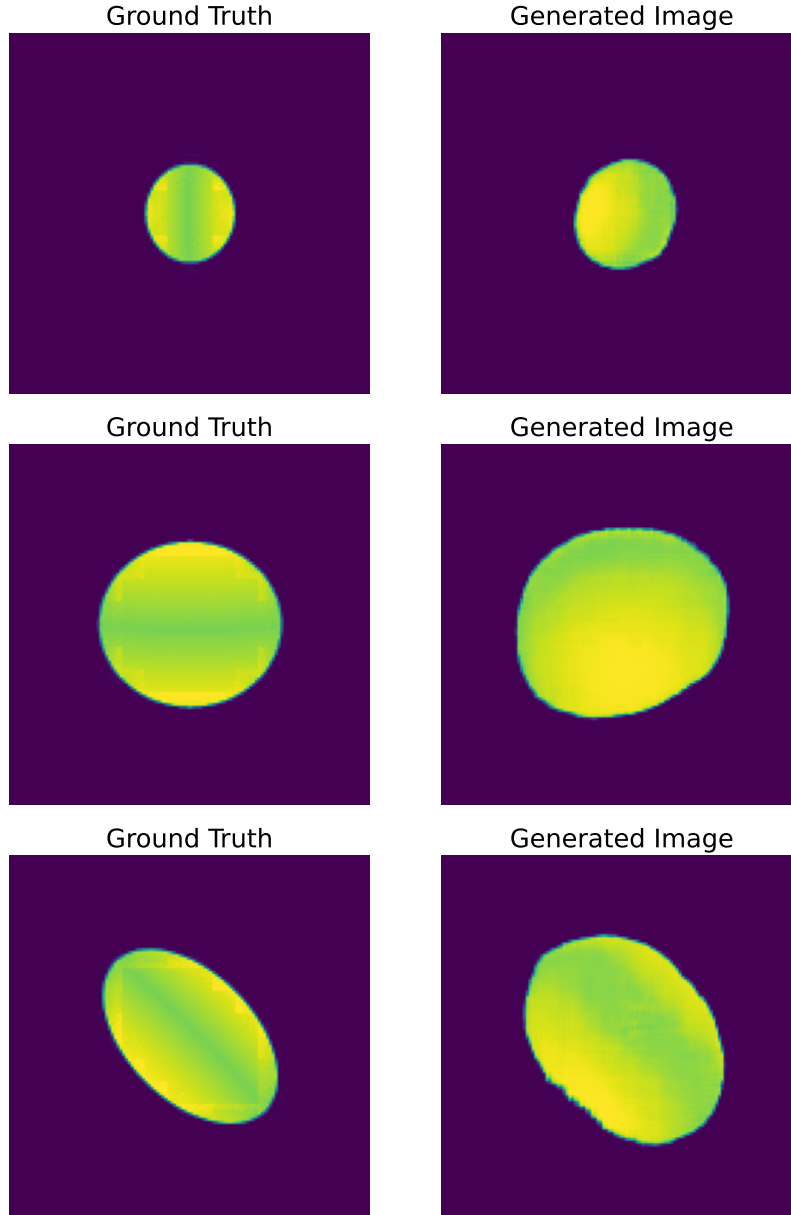


Figure 17: Simulated fast-rotating stars and the reconstructed images for the LST setup with 7 sub-mirrors. Here, the observation time is 3 hours, which is enough for a rough estimation of the size and the shape of the source.

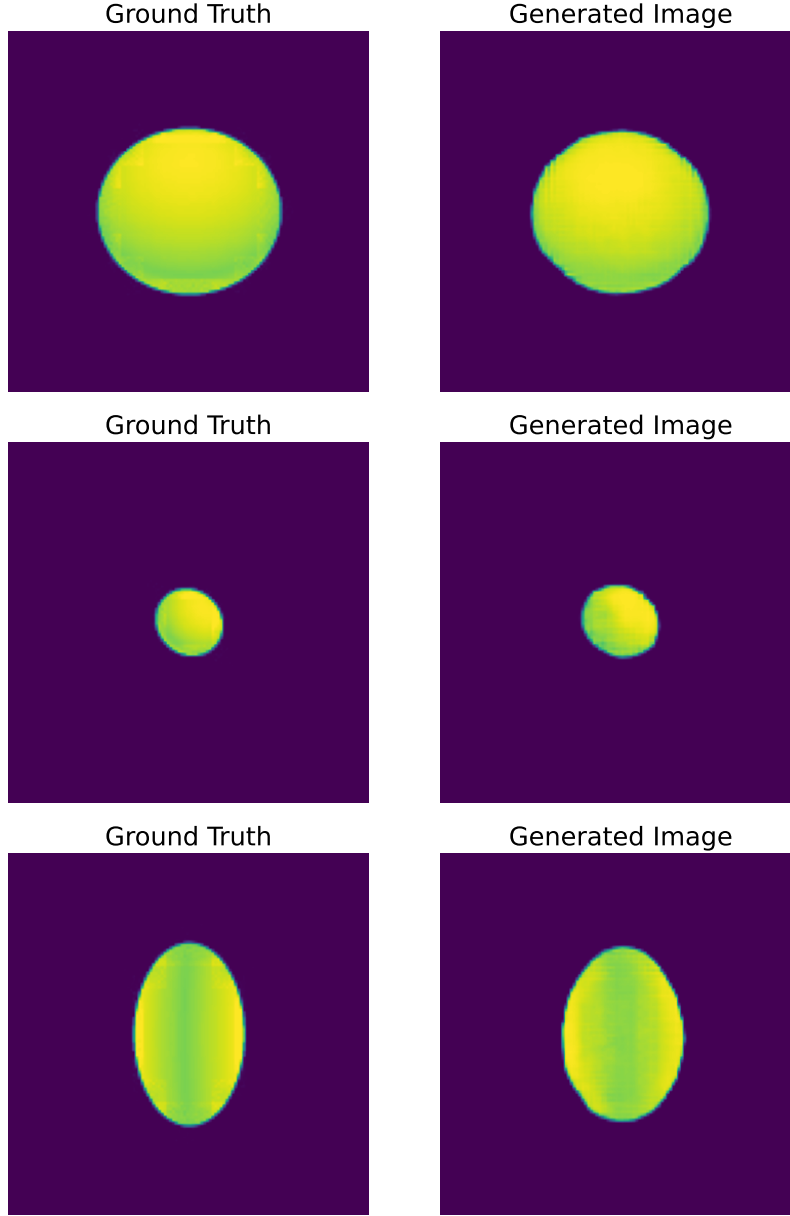


Figure 18: Simulated fast-rotating stars and the reconstructed images for the LST setup with 7 sub-mirrors. Here, the observation time is 6 hours, which gives great results, as in addition to the shape and the size, the brightness distribution is reconstructed with high accuracy.

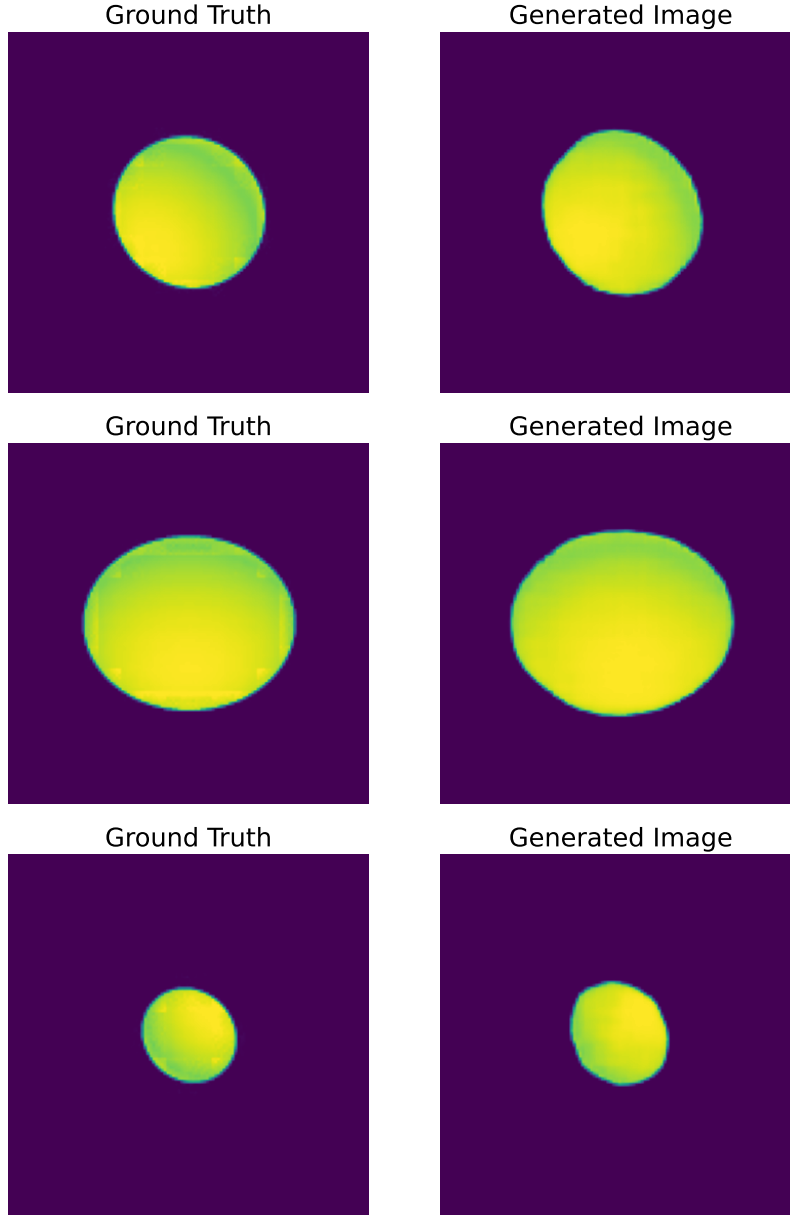


Figure 19: Simulated fast-rotating stars and the reconstructed images for the LST setup with 7 sub-mirrors. Here, the observation time is 9 hours, which gives great results.

5 Conclusion and outlook

The problem of phase retrieval in Intensity Interferometry is successfully approached with machine learning methods.

The main factor that determines the performance of the reconstruction is how much of the Fourier plane is sampled, which manifests as the number of available telescopes and the total observing time. Assuming a full night of observation, three telescopes allow a credible reconstruction of the shape and the size of the source. With four telescopes, we can reconstruct the brightness distribution with higher precision and reduce the observation time.

The cGAN presented in this work fails to reliably reconstruct sources with unknown sparse sampling. Further tests and improvements will be necessary if this is of interest.

The GAN is also applied to the specific layout of an LST at the northern CTA site. The performance of this layout depends on the total observation time, good results are achieved with six hours of observing. Nine hours do not seem necessary, because the result does not improve a lot.

More precise tests will be necessary to quantify the relationship between the amount of sparse sampling and the GAN performance, for example by varying the observation time only slightly on a given setup. It would also be interesting to explore different observatory layouts, for example, the southern CTA layout. As detector efficiencies are neglected in this work, it will be important to consider those and estimate the signal-to-noise ratio.

Conditional GANs are not the only method to create images from images. Other methods can be explored and compared. One can also adapt the GAN architecture. There are a lot of other loss functions, which can be implemented and compared. Even though various tests are needed to make the GAN even more robust and reliable, it is safe to say that machine learning is definitively a viable option for phase reconstruction in Intensity Interferometry.

References

- [1] C. Delgado et al., “Intensity interferometry with the MAGIC telescopes”, PoS **ICRC2021**, 693 (2021).
- [2] R. H. Brown and R. Twiss, “LXXIV. a new type of interferometer for use in radio astronomy”, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **45**, 663–682 (1954).
- [3] R. H. Brown, *The intensity interferometer: its application to astronomy* (Taylor & Francis; Halsted Press, London, New York, 1974), 184 pp.
- [4] V. A. Acciari, M. I. Bernardos, et al., “Optical intensity interferometry observations using the MAGIC imaging atmospheric cherenkov telescopes”, Monthly Notices of the Royal Astronomical Society **491**, 1540–1547 (2020).
- [5] D. Dravins, S. LeBohec, H. Jensen, and P. D. Nuñez, “Optical intensity interferometry with the cherenkov telescope array”, Astroparticle Physics, Seeing the High-Energy Universe with the Cherenkov Telescope Array - The Science Explored with the CTA **43**, 331–347 (2013).
- [6] J. Aleksić, S. Ansoldi, et al., “The major upgrade of the MAGIC telescopes, part II: a performance study using observations of the crab nebula”, Astroparticle Physics **72**, 76–94 (2016).
- [7] A. U. Abeysekara, W. Benbow, et al., “Demonstration of stellar intensity interferometry with the four VERITAS telescopes”, Nature Astronomy **4**, 1164–1169 (2020).
- [8] P. Saha, *The theory of intensity interferometry revisited*, Sept. 15, 2020.
- [9] I. J. Goodfellow, J. Pouget-Abadie, et al., *Generative adversarial networks*, June 10, 2014.
- [10] M. Mirza and S. Osindero, *Conditional generative adversarial nets*, Nov. 6, 2014.
- [11] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, *Image-to-image translation with conditional adversarial networks*, Nov. 26, 2018.
- [12] O. Ronneberger, P. Fischer, and T. Brox, *U-net: convolutional networks for biomedical image segmentation*, May 18, 2015.
- [13] M. Abadi, A. Agarwal, et al., “TensorFlow: large-scale machine learning on heterogeneous distributed systems”,

- [14] P. Virtanen, R. Gommers, et al., “SciPy 1.0: fundamental algorithms for scientific computing in python”, *Nature Methods* **17**, 261–272 (2020).
- [15] J. D. Hunter, “Matplotlib: a 2d graphics environment”, *Computing in Science & Engineering* **9**, 90–95 (2007).
- [16] K. P. Murphy, *Probabilistic machine learning: an introduction*, Adaptive computation and machine learning (The MIT Press, Cambridge, Massachusetts London, England, 2022), 826 pp.
- [17] S. J. D. Prince, *Understanding deep learning* (The MIT Press, Cambridge, Massachusetts, 2023).
- [18] M. Hayashida, K. Noda, et al., “The optical system for the large size telescope of the cherenkov telescope array”, (2015).