

CHRISTOPHER NEWPORT UNIVERSITY

CAPSTONE

Analysis of strong gravitational lenses using neural networks

Yemeen Ayub

supervised by
Dr. David GORE

Abstract

The study of gravitational lenses are paramount to understanding the shortcomings of the current model of cosmology. Unfortunately, we have only observed a small number of lensed systems. With this number expected to go up by several orders of magnitude in the next generation of sky surveys [2], a method to analyze these systems will have to be fast, efficient, and accurate. Using convolutional neural networks, we have made it possible to analyze about 100 images of strong gravitational lenses in a second on a single graphics processing unit. This method recovers the Einstein radius, complex x and y ellipticity, and the x and y position of the center of the lens. These values are comparable to other methods used to analyze strong gravitational lenses while being ten million times faster.

April 29, 2018

Introduction

Strong gravitational lenses are a rare occurrence, but allow us to probe deep into the how the universe truly functions. Through the study of strong gravitational lenses, physicists noticed that certain galaxies were displaying strong gravitational lensing, but the galaxies did not have nearly enough mass to display these effects. This is exactly what led to the idea that there must exist a matter that has mass, but is not visible to us. Physicists named this type of mass, dark matter. A large amount of data can be determined about dark matter using gravitational lenses. One case is when joined with measurements such as central velocity dispersion and stellar population analysis, strong lenses can be used determine the amount of dark matter that lies in the center of galaxy-scale halos [6].

Currently, not many strong lensing system are known, but as the amount of astronomical data being recorded goes up exponentially, the number of humans to analyze will not keep up. Methods that are currently employed require complex lensing codes, many data preparation steps, and maximum likelihood models that are computationally expensive [3]. These methods work currently, but are slow and will not scale efficiently. It is because of these reasons that a new method will be necessary to accurately determine information about strong lenses.

Neural networks will be key to reconciling the problem of not being able to compete with the amount of incoming data. Once trained, a neural network is capable of recovering parameter values almost instantaneously, require almost no specialization to use, and can be run on most personal computers. This means that after one neural network is created, it may be shared with the world, so that anyone can enjoy the accurate results. In cases such as observational astronomy, it will allow amateur astronomers to recover values of galaxies that they have pictured themselves. It is by using neural networks, that humans will be able to scale with the large amount of incoming data.

A neural network is simply a function that learns from it's mistakes, and seeks to minimize how wrong it is. This function can be trained to recover many features such as the next state in a dynamical system, the molecular structure of unknown compounds, and even determine if a patient has cancer from just the images of the x-rays. These highly flexible tools can be particularly useful for analyzing gravitational lenses of massive objects, which can be a meticulous job done by hand. Sometimes an accurate analysis of a gravitational lens can take a few weeks and require many different data preparation steps [3]. This project seeks to train a neural network on simulated gravitational lenses displaying singular isothermal ellipsoid density profile and predict five parameters: the Einstein radius (θ_e), the complex ellipticity (ϵ_x, ϵ_y), and the coordinates of the center of the lens (x, y). These characteristics of strong gravitational lenses are important for telling us about the spatial distribution of mass at kiloparsec and sub-kiloparsec levels, determining the kinematics and geometry of the universe, and for studying distant galaxies normally too far away to see [7]. Using this neural network, researchers will be able to better determine the innerworkings of the universe.

The particular neural network being used in this project is Ensai, a convolutional deep neural network that predicts the previously listed parameters[3]. This project's goal will be to take an untrained neural network with the same structure as Ensai, and train it with about half a million both simulated and observed images of gravitational lenses with the singular isothermal ellipsoid profile and compare the resulting parameters with what the network Ensai would calculate for these same images. These images are simulated by taking high quality images of real galaxies and using them as lensed background sources in order to simulate the effects of strong gravitational lensing. That means that images will suffer highly nonlinear image distortion that will sometimes quadruple the number of background galaxies seen. Each image will be fitted with statistical noise to simulate effects noticed in observed data. Because of the nature of this added noise, no two images will be exactly the same.

With recent use of neural networks to find gravitational lensing in large surveys of data [?], the automation of analysis of strong gravitational lenses will provide a faster pipeline for data to be taken, analyzed, and fitted to current models. The next generation of surveys will produce millions of images of the sky, and a fast and efficient model will be necessary to keep up with the volume of data. The improvement of automatic analysis of these systems will be crucial if the field of observational astronomy is to adapt.

Theory

Gravitational Lenses

Gravitational lenses are measured by the direct comparison of the path of light in a perfectly homogeneous universe to our current distribution of mass. That is, if the universe had an even layer of mass everywhere, then light would follow a much different path than it does through our very non-homogeneous universe. This means that the homogeneous universe is considered the control case when considering gravitational lenses.

To be more concise, gravitational lenses are governed by the lensing equation

$$\beta = \theta - \alpha(\theta) \quad (1)$$

where θ is the angular position in the plane of the sky. $\alpha(\theta)$ is the reduced deflection angle and can be given by

$$\alpha(\theta) = \frac{D_{ds}\hat{\alpha}(D_d\theta)}{D_s} \quad (2)$$

where $\hat{\alpha}$ is the true deflection angle[1]. This equation can be seen in effect in Figure 1

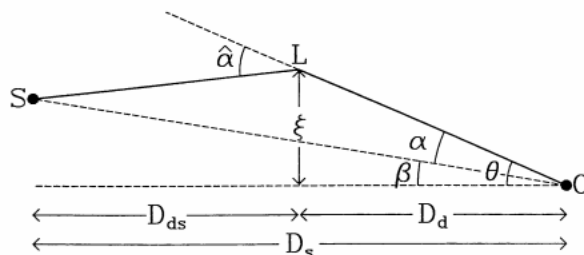


Figure 1: The lens equation visualized [1]

The true deflection angle $\hat{\alpha}$ requires several approximations in order to find an analytical solution. The first approximation is treating this object as a thin lens. This is justified as a thin lens system is characterized by the depth of the lens being much smaller than the distance to the source and the observer. Another fair approximation is based upon the fact that deflection angles are relatively small. This allows us to use the weak field approximation of General Relativity [1]. That is to say that we may ignore non-linear effects of General Relativity as they are generally pretty small in this case. This allows us to finally determine what $\hat{\alpha}$ is.

$$\hat{\alpha} = \nabla\psi/c^2 \quad (3)$$

where ψ is called the lensing potential and is given by

$$\psi(\theta) \equiv \frac{2}{c^2} \frac{D_{ds}}{D_d D_s} \int \Phi(D_d \theta, z) dz, \quad (4)$$

where c is the speed of light in a vacuum, D_{ds} is the lens-source angular-diameter distance, D_d is the observer-lens angular-diameter distance, D_s is the observer-source angular-diameter distance, and finally Φ is the gravitational potential of the matter placed along the line of sight on a thin plane placed between the observer and the background resources[4].

The gravitational potential, ψ , ends up being twice the 2D Newtonian potential and comes from solving for the 2D Poisson equation

$$\nabla^2 \psi(\theta) = 8\pi G \Sigma(\theta) \quad (5)$$

where $\Sigma(\theta)$ is the surface mass density. [1]

Gravitational lenses tend to be much more complex than optical lenses. This is exemplified by looking at the nature of ψ , the gravitational potential. This is a highly nonlinear function unlike the optical phenomenon which purely depends on the length of the lens.

Another key phenomenon of gravitational lenses is what is called an Einstein Ring. This effect can be observed when a light source is on the optical axis and is given by

$$\theta_e = \left(\frac{4GM}{c^2 D} \right)^{1/2} \quad (6)$$

where D is the effective lens distance and is given by,

$$D \equiv \frac{D_d D_s}{D_{ds}} \quad (7)$$

G is the gravitational constant, c is the speed of light in a vacuum, and finally M is a point mass [1]. This means that if a source is perfectly in line with the axis going through the center of the lens, then the apparent image to an observer will be wrapped in a circle of angular radius θ_e . This effect is rare and can be seen in Figure 2.

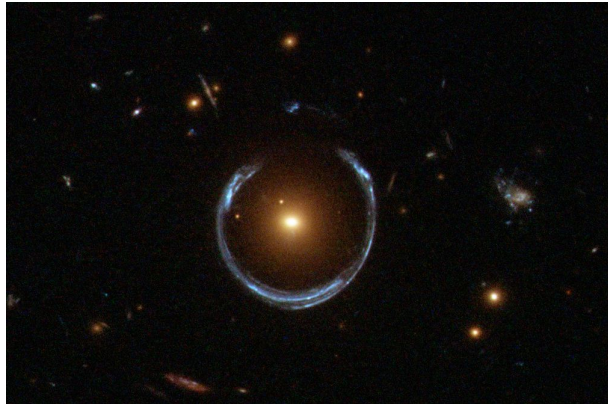


Figure 2: An Einstein Ring pictured by NASA

Another characteristic of a lens is its ellipticity or how ovoid an object is. This is important for defining the overall shape of the system. This can be measured using the complex ellipticity, which is given by

$$\epsilon_x = \frac{a^2 - b^2}{a^2 + b^2} e^{2i\phi}, \quad \epsilon_y = \frac{a - b}{a + b} e^{2i\phi}$$

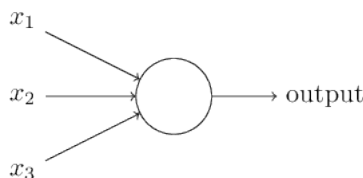
where a is the length of the semi-major axis, b is the length of the semi-minor axis and ϕ is the position angle. It's using this that we may easily characterize the traits of a gravitational lens.

Methods

The modelling of a gravitational lens can be very difficult as it is a highly nonlinear system. With our current methods, we may closely approximate the traits of a gravitational lens, but never truly model. Researchers have instead opted to model complex system of this nature using neural networks, which in itself is a highly nonlinear system.

Neural Networks

Just like the human brain, neural networks are comprised of neurons. A neuron is capable of taking in any amount of inputs and either returning a value. The neuron that either returns unity or a zero is called a perceptron. Deep learning neural networks are comprised of what are called sigmoid neurons, which means that instead of returning a discrete one or a zero, it can return a number anywhere between these two values. The neuron determines the value it will return by giving a certain weight to each input and determining the value it will return depending on the neuron. The perceptron takes the weight of all the inputs values and adds them until they reach a certain threshold value, there the perceptron flips the value it outputs to a one. This can be pictured in 3.



$$\text{output} = \begin{cases} 0 & \text{if } \sum_j w_j x_j \leq \text{threshold} \\ 1 & \text{if } \sum_j w_j x_j > \text{threshold} \end{cases}$$

Figure 3: A perceptron with three inputs [5]

These Neurons can be layered indefinitely, with each output feeding into the next layer of neurons. Where the final layer will then give the output that is favorable. In this case, the input is an image of a galaxy lens, and the output is the five parameters of the gravitational lens. The input of the image is split into each pixel, where a the pixel value is given by a zero if it is white, and a 1 if it is black. A value between that is assigned for any given shade the pixel may have. With that we feed in the pixels and receive an output value a . In order to teach the neural network, function to determine how wrong it's outputs are, a cost function must be determined.

$$C(a, y) = \frac{\sum_n |y_n - a_n|}{n} \quad (8)$$

where a is the neural network's output value, y is the expected value, and n is the number of training inputs. The cost function is to determine how “wrong” the neural network is. The best way to teach our neural network is to minimize this cost function. This means that the cost function must be pushed towards a value of 0 or as close to it as possible. Thankfully, there is a method for doing this called Stochastic Gradient Descent. It is given by

$$\Delta C \approx -\eta \|\nabla C\|^2 \quad (9)$$

where η is a small number called the learning parameter, and ∇C is the gradient of the cost function. This equation says that every small change in the cost function will be negative, thus the new value of C will always be less than the old one. Because our cost function is really just a function of the weights of the neural network, this will force the the weights of the network to change accordingly until $\Delta C \approx 0$.

The shortcomings of this method is that each layer of neurons tends to make training more computationally expensive. This was realized when neural networks were first theorized in the 1960's. With the exponential growth of computing power that's occurred since then, researchers may now enjoy the benefits of using a neural network from their home computer.

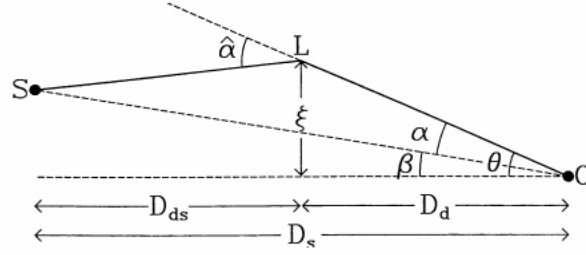


Figure 4: The lens equation visualized [1]

Training

About 60,000 images from the Galaxy Zoo project will be used as background images for the simulation [3]. The other background images will be simulated clumpy galaxies that are composed of 1-5 clumps [3]. These clumps will be generated using a Gaussian or Sérsic profile. These models are currently some of the best methods to model the distribution of mass in an elliptic galaxy. The simulated lens that will distort the galaxy will have parameters such as the ellipticity and the angle of the lens. These will be chosen at random from a flat distribution with a maximum ellipticity of 0.9. The coordinates of the the center of the lens are forced to cover close to the center of the image as it is necessary in order to demonstrate the full effects of the lensing. This forces the background lensing galaxy to be inside or close to the lens' caustic or no effects of strong gravitational lensing would be seen. Each image will consist of 192 x 192 pixels, with a pixel size of 0.04" [3].

Data Augmentation

At each step of gradient descent optimization, realistic effects are applied to the data. This allows for the prevention of over fitting to data or specializing the neural network to work especially well on the data set. Augmentation is also necessary so that the neural network will be better fit to actually assess real images of gravitational lenses. Furthermore, this will essentially ext

A Gaussian filter is applied with a randomly chosen root-mean-square value with a maximum of $0.1''$ to allow for a blurring effect. The image's intensity is then normalized using a factor of 100-1,000 for the image to photon count. This tied with gaussian noise creates a realistic noise effect

Data

The neural networks were put to the test on 1000 different images of gravitational images. These images were separated from the training data set in order to guarantee that the neural networks would not be superficially accurate. This is an important task when dealing with machine learning as it can guarantee very accurate results when the reality is that the neural network has adapted to that image's characteristics previously. The data was analyzed by evaluating each of the five parameter's root mean square errors (RMS). These errors were then compared to the errors of the previously implemented neural network, Ensai.

The neural network implemented in this project is called the AlexNet model. The set of data points can be pictured in figure 5. The RMS values were recorded to be $0.17''$, $0.24''$, $0.19''$, $0.13''$, $0.14''$ for the Einstein radius, x ellipticity, y ellipticity, center position on the x axis, and center position on the y axis, respectively. The set of data points for Ensai can be pictured in figure 6.

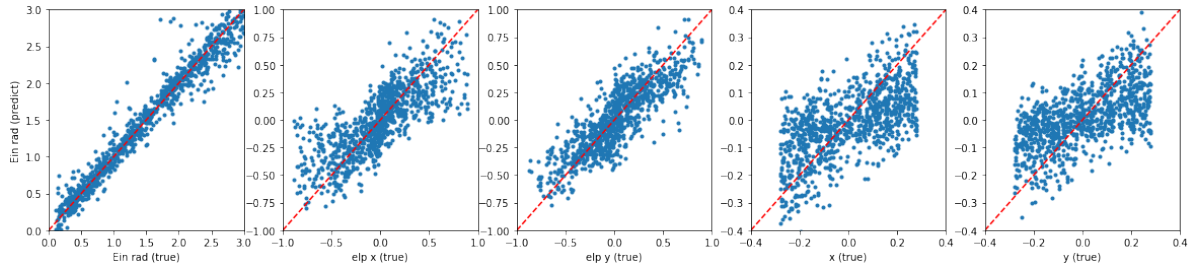


Figure 5: AlexNet model

The RMS values were recorded to be $0.05''$, $0.08''$, $0.10''$, $0.07''$, $0.07''$.

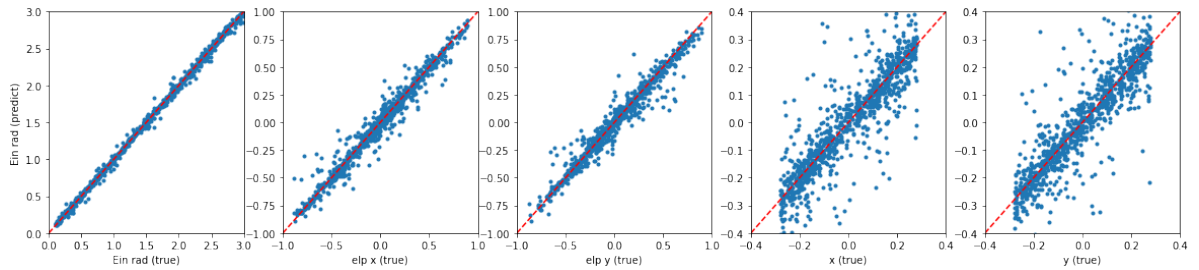


Figure 6: Ensai model

Discussions and Conclusions

Comparatively, AlexNet was consistently outperformed by Ensai. With an average difference in RMS value of $0.10''$. The closest AlexNet came to performing similarly to Ensai is with recovering the center position coordinates. This can be attributed to the fact that the center coordinates are much easier to determine with a strong lens that displays an Einstein ring. This kind of lensing system is considered to be rare as it depends on many variables. Most lensing systems will not display a perfect Einstein ring. This can be pictured in figure 7. The main reason that AlexNet gave much different results than Ensai is due to the difference in architecture of the neural network. Ensai's structure was much better suited for recovering ellipticity and Einstein radius, than AlexNet.

Future plans for this project include extending the pipeline of usage of this network. That is instead of only being able to provide images of gravitational lenses, it would be beneficial to be able to input any sky survey image, and return if there is a gravitational lens. If there does happen to be a lens, then the program would return the characteristics of the model. Another possible addition to this project would be to add the ability to input colored images. This would even further extend the pipeline.

Neural network have proven to be a very powerful tool. They have shown to be computationally inexpensive, quick, and accurate. It is using this method that future researchers will be able to quickly scan the night sky, and determine key features of the universe that otherwise would have gone unnoticed. It is this degree of automation that will be necessary in the coming years of when methods of telescopes with increasing accuracy are outputting millions of images to be used as data. It is only by employing neural networks will researchers be able to keep up with this costly demand.

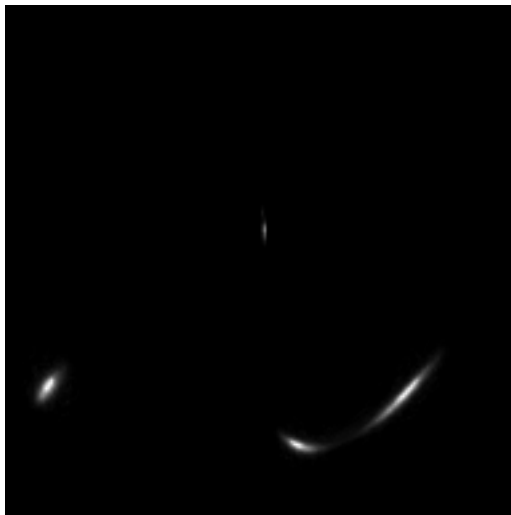


Figure 7: A simulated gravitational lens

References

- [1] R. D. Blandford and R. Narayan. Cosmological applications of gravitational lensing. *Annual Review of Astronomy and Astrophysics*, 30(1):311–358, 1992.

- [2] Thomas E. Collett. The population of galaxy–galaxy strong lenses in forthcoming optical imaging surveys. *The Astrophysical Journal*, 811(1):20, 2015.
- [3] Yashar D. Hezaveh, Laurence Perreault Levasseur, and Philip J. Marshall. Fast automated analysis of strong gravitational lenses with convolutional neural networks. *Nature*, 548(7669):555–557, Aug 2017. Letter.
- [4] Matteo Maturi, Sebastian Mizera, and Gregor Seidel. Multi-colour detection of gravitational arcs. *Astron. Astrophys.*, 567:A111, 2014.
- [5] Micheal Nielson. Neural networks and deep learning.
- [6] C. E. Petrillo, C. Tortora, S. Chatterjee, G. Vernardos, L. V. E. Koopmans, G. Verdoes Kleijn, N. R. Napolitano, G. Covone, P. Schneider, A. Grado, and J. McFarland. Finding strong gravitational lenses in the Kilo Degree Survey with Convolutional Neural Networks. , 472:1129–1150, November 2017.
- [7] T. Treu. Strong Lensing by Galaxies. , 48:87–125, September 2010.