HUNTING FOR DARK MATTER SUBSTRUCTURES IN STRONG LENSING WITH NEURAL NETWORKS

JOSHUA YAO-YU LIN (UIUC PHYSICS) NCTS DARK PHYSICS WORKSHOP 2020.1.9

Lin et al. 2020 (In prep)

About myself



- NTHU Physics (Undergrad 2012)
- NTU Physics (MS 2015)
- UIUC Physics (Ph.D. candidate 2016-present)

Research Interest:

- Dark Matter Substructures in Strong Lensing

- Cosmic Neutrinos

- Black Holes images, Radio radio interferometry, Machine Learning

Gravitational Lensing of Cosmic Neutrino Background

Lin & Holder (arXiv:1910.03550)

Joshua Yao-Yu Lin (University of Illinois at Urbana-Champaign)

In collaboration with Gilbert Holder

2020.1.9 NCTS Dark Physics Workshop



Lensing of CvB vs CMB

Similar:

Difference:

 Both (shall) have anisotropy and would be lensed by foreground gravitational potential





Thought Experiment!

Lensing of CvB vs CMB

Similar:

 Both (shall) have anisotropy and would be lensed by foreground gravitational potential



Image Credit:1) Planck 2) ESA/NASA, Hubble



Difference:

- Neutrinos from the early universe will be non-relativistic today (massive neutrinos)
- Larger angles of deflection
- Closer surface of last scattering (compared to the cosmic microwave background)
- Could form multiple lensed images [Strong gravitational lensing]

Cosmic Neutrino Last Scattering Surface



FIG. 2: The probability that a neutrino with mass m last scatters at a given comoving distance from us (the visibility function). Massive neutrinos travel more slowly than massless neutrinos so arrive here from much closer distances. Also shown is the last scattering surface of the cosmic microwave background, virtually indistinguishable from that of an $m_{\nu} = 10^{-4}$ eV neutrino.

$$v(a) = \frac{v_0}{\sqrt{a^2 + \frac{v_0^2}{c^2}(1 - a^2)}}$$

$$D_S(v_0) = \int_{a_s}^1 \frac{da}{a^2 H(a)} v(a)$$

Dodelson & Vesterinen (PRL, 2009)

Cosmic Neutrino Last Scattering Surface



Strong lensing of cosmic neutrino

Lens equation



 Angle of deflection for cosmic neutrino (post-Newtonian)

$$\alpha(R) = \frac{4GM(R)}{Rc^2} \frac{c^2 + v_{lens}^2}{2v_{lens}^2}$$

Slides credit: Sherry Suyu

Einstein Radius: function of neutrino velocity/ Distance to Lens



$$\frac{dP}{dp_0} = \frac{2}{3\zeta(3)k_B^3 T_v^3} \frac{p_0^2/c^3}{e^{p_0/k_B T_v} + 1}$$

Neutrinos momentum distribution

$$\theta_{E}^{\text{SIS}} = \frac{4\pi\sigma_{v}^{2}}{c^{2}} \left(\frac{c^{2} + v_{lens}^{2}}{2v_{lens}^{2}}\right) \frac{D_{LS}(v_{0}, D_{L})}{D_{S}(v_{0})}$$

SIS lens model: $\sigma v = 1000$ km/s

Mass eigenstates splitting via gravitational potential



Neutrino source (flavor)

m, m, m, Normal



Gravitational Lens: Spectrometer



Mass eigenstates splitting

Mass eigenstates

Spectrometer: Stern–Gerlach experiment



Mass eigenstates splitting via gravitational potential



(a) Neutrino flux for mass eigenstates



(c) Neutrino flux for muon neutrinos



(b) Neutrino flux for electron neutrinos



(d) Neutrino flux for tau neutrinos

Discussion & Summary

- Strong lensing of CvB could be the largest Stern–Gerlach experiment in our universe
- Neutrino oscillations are not relevant in this case, as the mass eigenstates get dispersed in angular space
- Time evolution of the halo could also be probed
- Interesting, with extremely rich source of information: strong lensing, neutrino mass, quantum properties



Lin & Holder (arXiv:1910.03550)



University of Illinois

Image credit: Aram Grigoryan/Getty Images

HUNTING FOR DARK MATTER SUBSTRUCTURES IN STRONG LENSING WITH NEURAL NETWORKS

JOSHUA YAO-YU LIN (UIUC PHYSICS) NCTS DARK PHYSICS WORKSHOP 2020.1.9

Lin et al. 2020 (In prep)

PEOPLE I WORK WITH LOCAL GROUP









Hang Yu [UIUC] Warren Morningstar [Stanford] Jian Peng [CS@UIUC]

Gil Holder [UIUC]

The multiple components that compose our universe

Current composition (as the fractions evolve with time)



Three different types of neutrinos comprise at least 0.1%, the cosmic background radiation makes up 0.01%, and black holes comprise at least 0.005%.

David Spergel (Science, 2015)

Dark Matter

- 84% of the matter is Dark(DM)
- DM interacts through gravity.
- Further DM interactions unobserved so far.
 Such couplings must be very weak, much weaker than weak interactions.

N-BODY SIMULATION: STUDY THE PROPERTY OF DARK MATTER



Max-Planck-Institut für Astrophysik (2005)

N-Body Simulation on Cold Dark Matter

In various cosmological N-body simulation, the A Cold Dark Matter (ACDM) model preform well especially on the large scale structure. (e.g. Millennium Run 2005)!

Max-Planck-Institut für Astrophysik (2005)

Controversies on Galactic Scale

-- Comparison of Numerical Simulation & Observational Data

Missing Satellite problem





D. H. Weinberg et al. (2013) (arXiv:1306.0913) 21

Missing Satellites Problem



CDM Simulation (Mayer and Kazantzidis)

Satellite galaxies of the Milky Way (Observation)

Strong lensing: Natural place to probe dark matter substructures



Strong lensing



Lensing basic

Lens equation



$$\boldsymbol{\eta} = \frac{D_{\mathrm{s}}}{D_{\mathrm{d}}} \boldsymbol{\xi} - D_{\mathrm{ds}} \hat{\boldsymbol{\alpha}}(\boldsymbol{\xi})$$

In terms of angular coord.:

$$oldsymbol{\eta} = D_{\mathrm{s}}oldsymbol{eta}$$

 $oldsymbol{\xi} = D_{\mathrm{d}}oldsymbol{ heta}$

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta})$$

where

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{D_{\mathrm{ds}}}{D_{\mathrm{s}}} \boldsymbol{\hat{\alpha}}(D_{\mathrm{d}}\boldsymbol{\theta})$$

Slides credit: Sherry Suyu

Lensing basic



Lens Image



Figure from Narayan & Bartelmann (1995)

Subhalo detection



 $ec{lpha}_{lens} = ec{lpha}_{smooth} + \sum_{i=0}^N ec{lpha}_{i,subhalo}$

Figure 2. A simple sketch of the method we used to create our mock data; subhaloes and line-of-sight haloes are placed so that their lensing effect lies in the same projected position on the plane of the main lens; the grey region gives an example of the line-of-sight volume that is taken into account.

Despali et al. (1710.05029)

Strong lensing with substructure as perturber



Lensing without perturber



Lensing with perturber



Lensing with (large) massive perturber

400

200

0

600

Subhalo Hidden in ALMA Gravitational Lens Image

Hezaveh (2016) ArXiv:1601.01388



Figure 6. Top left: the sky emission model in band 6 for the best-fit smooth lens parameters for the SDP.81 data. Top middle: the same for the perturbed model. Top right: the difference between the two models. The bottom panels show the same for band 7. The bright feature in the difference plots is mainly caused by the astrometric anomaly of the arc. In each row, the images have been scaled to the peak flux of the smooth model.

Hezaveh et al., (2016)

LOOKING INTO THE FUTURE: New Lenses

For future surveys we find that, assuming Poisson limited lens galaxy subtraction, searches of the DES, LSST, and Euclid data sets should discover 2400, 120000, and 170000 galaxy-galaxy strong lenses, respectively

Collett, ApJ. 2015



Slides from Laurence Perreault Levasseur

Can AI (deep learning) help?



Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

Image credit: Nvidia

Brief Intro to deep learning: MNIST dataset





Credit: http://yann.lecun.com/exdb/mnist/

Brief Intro to deep learning: Deep Neural Networks



Credit:LeNet by Yann LeCun (1998)

Examples of handwritten digits that can be recognized correctly the first time they are seen

Slides from Geoff Hinton (Coursera, 2012)

DenseNet and ResNet





"Deep Residual Learning for Image Recognition" Kaiming He et al (arXiv:1512.03385)
Loss Surface w Skip Connection



Li et al. NIPS 2017 (ArXiv: 1712.09913)

CONVOLUTIONAL NEURAL NETWORKS: PREVIOUSLY USED TO FIND LENSES (CLASSIFICATION)

THEY CAN BE TRAINED TO CLASSIFY IMAGES: TWO CLASSES: LENSES VS. NON-LENSES



CMU DeepLens: Deep Learning For Automatic Image-based Galaxy-Galaxy Strong Lens Finding

François Lanusse,¹ Quanbin Ma,² Nan Li,^{3,4} Thomas E. Collett,⁵ Chun-Liang Li,²

Siamak Ravanbakhsh,² Rachel Mandelbaum¹ and Barnabás Póczos²

School of Computer Science, Carnegie Mellon University, Philaburgh, PA 15213, USA

³ High Energy Physics Division, Argonne National Laboratory, Lemont, IL 60439, USA

Finding Strong Gravitational Lenses in the Kilo Degree Survey with Convolutional Neural Networks

C. E. Petrillo^{1*}, C. Tortora¹, S. Chatterjee¹, G. Vernardos¹, L. V. E. Koopmans¹, G. Verdoes Kleijn¹, N. R. Napolitano², G. Covone³, P. Schneider⁴, A. Grado², J. McFarland¹

¹ Kapteyn Astronomical Institute, University of Groningen, Posthus 800, 9700 AV, Groningen, The Netherlands
² ENAF - Osservatorio Astronomico di Capodimante, Salita Moiariello, 16, 80131 Napoli, Italy
³ Dipartimento di Scienze Fisiche, Università di Napoli Federico II, Compl. Univ. Monte S. Angelo, 80126 Napoli, Italy
⁴ Argelander-Institut für Astronomie, Auf dem Higel 71, D-53121 Bonn, Germany

¹McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15813, USA

⁶Department of Astronomy & Astrophysics, The University of Chicage, 5640 South Ellis Avenue, Chicage, IL 60637, USA ⁵Institute of Cosmology and Gravitation, University of Portsmouth, Burnalty R4, Portsmouth, PO1 3FX, UK

Fast automated analysis of strong gravitational lenses with convolutional neural networks

Yashar D. Hezaveh^{1,2}*, Laurence Perreault Levasseur^{1,2}* & Philip J. Marshall^{1,2}

Quantifying image distortions caused by strong gravitational lensing—the formation of multiple images of distant sources due to the deflection of their light by the gravity of intervening structures-and estimating the corresponding matter distribution of these structures (the 'gravitational lens') has primarily been performed using maximum likelihood modelling of observations. This procedure is typically time- and resource-consuming, requiring sophisticated lensing codes, several data preparation steps, and finding the maximum likelihood model parameters in a computationally expensive process with downhill optimizers¹. Accurate analysis of a single gravitational lens can take up to a few weeks and requires expert knowledge of the physical processes and methods involved. Tens of thousands of new lenses are expected to be discovered with the upcoming generation of ground and space surveys^{2,3}. Here we report the use of deep convolutional neural networks to estimate lensing parameters in an extremely fast and automated way, circumventing the difficulties that are faced by maximum likelihood methods. We also show that the removal of lens light can be made fast and automated using independent component analysis4 of multi-filter imaging data. Our networks can recover the parameters of the 'singular isothermal ellipsoid' density profile5, which is commonly used to model strong lensing systems, with an accuracy comparable to the uncertainties of sophisticated models but about ten million times faster: 100 systems in approximately one second on a single graphics processing unit. These networks can provide a way for non-experts to obtain estimates of lensing parameters for large samples of data.

deep learning, convolutional neural networks (Methods) have been shown to excel at many image recognition and classification tasks⁶. This makes them a particularly promising tool for the analysis of gravitational lenses. Recently, these networks have been used to search for gravitational lenses in large volumes of telescope data⁷⁻⁹ and to simulate weakly lensed galaxy images¹⁰. Here we show that these networks can also be used for data analysis and parameter estimation.

We train four networks, Inception-v411, AlexNet12, OverFeat13 and a network of our own design, to analyse strongly lensed systems, by removing their final classification layer and interpreting the outputs of the last fully connected layer as a prediction for lensing parameters, with all weights initialized at random. We train the networks to predict the five parameters of the singular isothermal ellipsoid profile: the Einstein radius, the complex ellipticity and the coordinates of the centre of the lens. We use a squared-difference cost function, averaged over the five parameters. Although in many situations in machine learning collecting sufficiently large training sets is one of the main challenges, here it is possible to simulate the training data extremely fast. We train the networks on half a million simulated strong lensing systems. The lensed background sources are composed of three equal sets of images: the first and second comprise real galaxy images from the Galaxy Zoo14 machine learning challenge and high-quality images from the GREAT3 training data¹⁵, and the third set is composed of simulated clumpy galaxies with Sérsic and Gaussian clump profiles. The position of the background galaxy in the source plane is chosen randomly for each sample, but limited to regions where strong lensing occurs, that is, inside or on the caustics.

Hezaveh et al. (Nature, 2017)

PRODUCING THE TRAINING DATA



Fast automated analysis of strong gravitational lenses with convolutional neural networks



EXAMPLES OF SIMULATED DATA

500,000 simulated data

 "label" SIE model with 5 parameters: Einstein radius[θ_E (arcsec)], x, y, e_x, e_y

| Table 1 Errors of the individual and combined networks | | | | | |
|--|---------------------------|-----------------------|-----------------|------------|------------|
| Network | $\theta_{\rm E}$ (arcsec) | $\varepsilon_{\rm X}$ | ε_y | x (arcsec) | y (arcsec) |
| Inception-v4 ¹¹ | 0.03 | 0.04 | 0.05 | 0.06 | 0.06 |
| AlexNet ¹² | 0.03 | 0.04 | 0.04 | 0.05 | 0.06 |
| OverFeat ¹³ | 0.04 | 0.05 | 0.05 | 0.06 | 0.06 |
| Our network | 0.03 | 0.05 | 0.06 | 0.05 | 0.05 |
| Combined network | 0.02 | 0.04 | 0.04 | 0.04 | 0.04 |

The columns present the 68% errors for the Einstein radius (θ_E), the x and y components of complex ellipticity (ε_x and ε_y), and the coordinates of the lensing galaxy (x and y) for each individual network and the combined network. The angular parameters (θ_E , x and y) are given in units of arcseconds.

4 CNNs gives pretty good results!

Hezaveh et al. (Nature, 2017)

Could we detect dark matter subhalos in Strong lensing with Neural networks?

- Simulated data could help us understand the problem
- strong gravitational lensing are insufficient (≤ O(10)) for typical size of the training set for deep learning
- Simulated data would be good for supervised learning since we know the ground truth

Deep learning setup

20000 simulated data as training set



- 20000 simulated data [images and subhalo ground truth]as training set
- 2000 "DIFFERENT" data [images] as test set
- Simulation with SIE (marco lens) contains 0-5 subhalos(perturbers)
- Loss function: Binary Cross Entropy Loss (of subhalo probability map) $BCE = t_i \log(p_i) + (1 - t_i) \log(1 - p_i)$

PYTORCH

- Adam Optimizer, learning rate = 1e-4
- NN model: DenseNet (~53 layers)
- Nvidia GPU: 1080Ti

DenseNet



DenseNet architecture (121 layers)

Gao Huang et al., ArXiv:1608.06993



Let's check how NN is doing



Prediction: subhalo detected!



Prediction: subhalo detected!



Prediction: multiple subhalos!



Prediction: multiple subhalos!



"Failed" good examples: Can't see in the dark



"Failed" good examples: Can't see in the dark



Prediction: No subhalo



0.0

Prediction: No subhalo



"Rejection" of subhalo(s) around the arc



"Rejection" of subhalo(s) around the arc



Summary for Sim I

- Deep learning shows some promising result in dark matter substructures detection in lensing.
- "rejections" for no subhalos around the strong lensing arc.
- "detections" and "regression" for subhalos around the strong lensing arc.
- More realistic simulation needed

Simulation II

- Simulation II: More realistic, Evillens Based
- Sources
- Power law elliptical model
- Subhalos Pseudo-Jaffe Profile, Cumulative mass function, and radial distribution function from Aquarius Simulation
- Angular resolution: 0.02"
- 500,000/ 100,000 images as train/test set

Simulation II



- Perturbations are small
- Most of the subhalos are not detectable (far away from lensed source)

Detectable subhalos

- Subhalos near lensing arc
- Einstein radius of subhalo overlap with lensed source (>max pixel/15)
- We treat all the subhalos (with different masses) with the same target probability density



Prediction: subhalo detected!



Prediction: subhalo detected!



Multiple subhalos detected



Discussion

- Neural network are able to find subhalos in a more realistic simulation
- Unlabeled subhalos could be detected if close enough to the lens
- Some summary statistics would be needed





• What about Low mass subhalos?

CDM Simulation (Mayer and Kazantzidis)

Lots of low mass (< 10^8 M_sun) subhalos contributes!



CDM vs SIDM



Lots of low mass subhalos contributes in CDM!

Image credit: Mark Vogelsberger

Power spectrum of DM models



Figure 3. Top Left: redshift dependence of the convergence power spectrum for the CDM simulation. Top Right: redshift dependence of the convergence power spectrum for the ETHOS4 simulation. Bottom Left: mass dependence of the convergence power spectrum for the CDM simulation. Bottom Right: mass dependence of the convergence power spectrum for the ETHOS4 simulation. Note that the y-axis is the same for a given row but differs between rows. The wavenumbers k are in comoving coordinates. *As discussed in the text, the z = 0 power spectra are computed using the subhalo catalog at z = 0 but the distance between the observer and the lens D_{ol} is fixed to be the same as for a lens at z = 0.5 because Σ_{crit} diverges as $z \to 0$.

Díaz Rivero et al. (ArXiv:1809.00004)

Statistical detection of dark matter power spectrum



Smooth Lens

with Gaussian Random Field

Difference

Noise



Likelihood function

(9)

$$\mathcal{L}(\boldsymbol{O}_{\text{obs}}, \boldsymbol{p}) = \int d^{n} N d^{2n} \alpha P(\boldsymbol{N}) P(\boldsymbol{\alpha})$$
$$\delta \left[\boldsymbol{O}_{m}(\boldsymbol{p}) + \frac{\partial O}{\partial \alpha} \Delta \boldsymbol{\alpha} + \boldsymbol{N} - \boldsymbol{O}_{\text{obs}} \right] P_{p}(\boldsymbol{p}) \qquad (6)$$

$$\mathcal{L}(C_{\alpha}) = (|C_{N}| |C_{\alpha}| |C_{p}| |M|)^{-1/2} e^{\frac{1}{2}B^{\mathsf{T}} M B} \\ e^{-\frac{1}{2}(\Delta \mathbf{O}^{\mathsf{T}} C_{N}^{-1} \Delta \mathbf{O} + \mathbf{p}_{0} C_{p}^{-1} \mathbf{p}_{0})}$$



FIG. 2.— Joint-likelihood of noise and the amplitude of the power spectrum, mapped by evaluating Equation (9) using mock observations described in §4, lensed by a density field which includes substructure with a flat power spectrum. The dashed lines show the true values which were used in the mock observation. The input amplitude of the power spectrum is successfully recovered, with little if any degeneracy between instrumental noise and substructure fluctuations.





Hezaveh et al. (2016)

Preliminary result with NN





1.0 0.9999997763464862 0.329908 0.33957374




Discussion

- Maximum likelihood could serve as a tool to probe power spectrum of dark matter substructures in strong lensing system
- ML also could serve as a interesting tool for DMS power spectrum

Thank you!



Image Credit: Hubble/STScI & NASA

Prediction: unlabeled subhalo detected!



Detection: Complicated Lens



No detection



Fail to detect



False positive: Degeneracy

lensing image smooth model source galaxy 25 -75 -175 -200 -200 -



Degeneracy



True positive

