University of Groningen

Automating the detection of strong gravitational lenses in large-scale surveys using deep learning

Nagam, Bharath Chowdhary

DOI:
10.33612/diss.1187888207

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2025

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Nagam, B. C. (2025). *Automating the detection of strong gravitational lenses in large-scale surveys using deep learning*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen. https://doi.org/10.33612/diss.1187888207

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 19-04-2025

6

CONCLUSION

*In the everlasting cosmic dance, there is no beginning nor an end (**conclusion**). Birth and death are mere illusions. From matter and energy to the Universe itself, forms keep changing from one to another. Change is the only constant in the Universe. Even time exists at the mercy of change. Without change, there is no time. We are always **here, now** in the infinite present, witnessing this everlasting change.*

ABSTRACT OF THIS THESIS

Automated morphological classification of strong gravitational lenses in vast astronomical imaging datasets presents a significant challenge due to their inherent rarity and similarity to some other cosmic objects. This can result in a high rate of false positives, easily exceeding a hundred non-lenses for each identified lens candidate. The advent of missions like Euclid, that survey large areas of the sky at high angular resolution and large depth, has underscored the impracticality of human visual inspection for datasets encompassing, in some cases, up to 10 billion objects. Convolutional Neural Networks (CNNs) have emerged as a highly effective tool for automating the discovery of strong lenses, as demonstrated in previous lens-finding challenges (Petrillo et al., 2017; Metcalf et al., 2019). In this thesis, diverse methodologies for identifying strong gravitational lenses and rejecting false positives in large-scale surveys were explored, going beyond the traditional reliance solely on morphological classification, i.e., P value. Below, a brief summary of the findings and an outline of potential future directions for research and application are provided.

6.1 DENSELENS – GOING BEYOND STATE-OF-THE-ART RESNETS

In Chapter 2, the aim was to explore ways to reduce the false positives present in the current state of the art CNN architectures. A DenseNet architecture was introduced for strong gravitational lens detection, comparing its performance to the then state-of-the-art

ResNet models developed by Petrillo et al. (2017, 2019a,b) and Li et al. (2020, 2021). It is shown that the DenseNet-121 ensemble outperforms ResNet architectures in classifying lenses and non-lenses, particularly in reducing false positives, which are currently the main bottleneck in automatically classifying strong gravitational lenses in large surveys. At a false positive rate of 10^{-3}, DenseNet-121 achieves the highest true positive rate of 0.68 while requiring fewer parameters. A novel Information Content (IC) metric was introduced for rank-ordering lens candidates, which scales with the number of spatial resolution elements above a brightness threshold and the ratio of the Einstein radius over source effective radius, with easier-to-recognize lenses having higher IC values. A pipeline ensemble model combining classification and regression CNNs was developed to filter and rank-order candidates based on IC values, further reducing the need for visual inspection in large surveys. There is a demonstrated good correlation between predicted and actual IC values, and improved performance of DenseNet across the range of Einstein radii distributions is shown when including both metrics, over using only the classification prediction value (P). Despite the large reduction in false positives, they still largely outnumber the true lenses for any reasonable true-positive rate and hence additional improvements of the network are explored in Chapter 3.

6.2 U-DENSELENS – COMBINING MULTIPLE METRICS AND APPLICATIONS TO KIDS DATA

In Chapter 3, the aim was to design additional metrics to further aid in the reduction of false positives and pave the way for automation of strong lens detection. A U-Net segmentation algorithm was introduced, specifically built to separate the image part containing the lensed source from the rest of the image (i.e., lens galaxy and field), to complement the Denselens classifier, further enhancing strong gravitational lens detection in large surveys. The U-Net algorithm, integrated with classification scores (P) and Information Content (IC) metrics, significantly improves the purity of lens candidates. Analysis of a mock dataset demonstrates that filtering candidates with a segmentation score (n_s=0) can eliminate 77% of false positives while retaining 86% of true lenses, in a balanced dataset. Given that real datasets from e.g., KiDS and Euclid are unbalanced, with of order a thousand non-lens galaxies per one lens galaxy, the reduction in false positives is in practice a factor of about four. A decision tree was implemented for the Bright Galaxy (BG) sample from KiDS, using the P, IC, and n_s metrics to identify the primary contributor in the decision-making process. This approach led to the discovery of fourteen new strong lensing candidates. The method's effectiveness was further confirmed on the Luminous Red Galaxy (LRG) sample from KiDS, reducing false positives by 25%. While the classifier is optimized for KiDS r-band data, it is expected that this decision tree framework will prove robust for automating strong lens detection in future large-scale surveys, such as Euclid, as shown in Metcalf et al. (2019).

6.3 Denoising Diffusion GAN – increasing the training set of false positives

In Chapter 4, with the aim in mind to improve the training of the networks to reject non-lenses, I have applied Denoising Diffusion Generative Adversarial Networks (DDGAN) to create very large sets of simulated images of strong lens ’impostors’, utilizing non-lens candidates with high CNN scores. This approach shows promise in reducing false positives when used in conjunction with the Denselens algorithm. The DDGAN model demonstrates the ability to conditionally generate samples that are visually indistinguishable from the input distribution, offering a potential tool for data augmentation in future CNN training. While our DDGAN-generated data visually resembles real observations, I did find that training CNNs exclusively on simulated data led to performance deterioration when tested on the KiDS-based dataset or using them also for creating lenses. Notably, I have shown that a balanced combination of real and simulated samples in the training set can yield CNN performance comparable to using only real data. Our experiments reveal that the Denselens CNN ensemble model can achieve robust performance with as little as 25% real data in the training set, which is particularly promising for future large-scale surveys where annotated real data may be scarce. These findings underscore the potential of generative models in enhancing gravitational lens detection techniques, while also highlighting areas for future improvement and optimization.

6.4 Application to Euclid data – the first ML-detected strong lenses

To find the first strong lenses in the Euclid survey, in Chapter 5, I applied `Denselens` and `U-Denselens` algorithms to Euclid Early Release Observations (ERO) data. Our analysis, focusing on the Perseus field and extending to 16 other ERO fields, demonstrates the performance of `U-Denselens`. In the Perseus field, where human classifiers identified 68 potential lens candidates, `Denselens` ($P_{\mathrm{THRESH}} = 0.23$) detected 34 true positives but produced 558 false positives. In contrast, `U-Denselens` ($P_{\mathrm{THRESH}} = 0.18$) identified 27 true positives with only 331 false positives, reducing the false positive rate by 41%, while maintaining high detection efficiency. Across all ERO fields, I discovered 46 strong lens candidates, including 12 high-confidence detections with clear lensing features. `U-Denselens` consistently outperformed `Denselens`, identifying nearly all candidates detected by `Denselens` while substantially reducing false positives. Although the contamination rate (FP/TP ratio) is still considerable, the contamination rate is significantly lower for Euclid data (16) compared to the KiDS data (40), even for a much lower P-value threshold. This reflects that the networks trained on lower-quality data work extremely well on Euclid data and, due to its superior data quality, yields a much lower contamination rate. For a total area of 8.12 square degrees, I identified 3 highly probable lenses at a rate of 3.67 false positives for every true positive found. Thus, from the entire Euclid survey covering about 15,000 square degrees, I expect identifying at least already 5,500 highly probable strong lenses with this provisionally trained network, at the expense of approximately 20,000 false positives. These results highlight the potential of `U-Denselens` to enhance through thorough training and tuning the efficiency of gravitational lens searches in large-scale surveys, offering a promising approach to streamline the identification process in extensive astronomical datasets.

FUTURE WORK

TRAINING REFINEMENT

Non-lenses: Our work has identified the most common types of false positives arising from the U-Denselens network. Moving forward, I propose retraining our networks with the initial false positives from Euclid data, and also by using real Euclid data to create mock lenses on which the networks are trained. As research with DDGANs progresses and simulations improve in matching real data distributions, I aim to train our networks using simulated non-lenses generated by the DDGAN algorithm.

Lenses: An alternative approach involves improving the realism of creating mock lenses and training our networks exclusively with lenses that have a high Information Content (IC) value. I remove from the training sample the positives with such low IC values that they are not easily identifiable as lenses by human classifiers. Subsequently, I will retrain the entire ensemble network using only the remaining, clearly human-identifiable positive samples with high IC values, thereby enhancing the robustness of the model and minimizing the impact of less reliable data.

INTEGRATION WITH EUCLID

In the coming years, I plan to apply our currently-optimal networks to the entire Euclid survey dataset. Our goal is to incorporate these networks into the Euclid pipeline, enhancing its capability to recover strong lenses efficiently and accurately.

ADVANCING TO GPT-LIKE REASONING

In the long-term, I envision Generative Pre-trained Transformers (GPT; Brown et al., 2020) as the future of astronomical analysis. These models hold immense potential for modeling and classification tasks in astronomy. Imagine a single, comprehensive model capable of not only providing modeling results and classifying lens candidates but also articulating its reasoning behind classifications as lens, non-lens, or potential lens. Currently, result interpretation and analysis are limited to those working directly with the models. GPT-like models could democratize this process, allowing humans to interact with the model, fostering a deeper understanding of lens candidates and accelerating scientific discovery in the field.