FERMILAB-POSTER-23-212-STUDENT

Benchmarking Simulated Gravitational Lensing Classification with Neural Nets

Antonella Macoretta, Johns Hopkins University – SIST Intern | Brian Nord, Alex Drlica-Wagner

Background

When a massive object curves space time, the path of light passing through it is bent, causing gravitational lensing. In **strong lensing (SL)**, galaxies are the massive foreground object.

We lack good, non-manual ways to find galaxy-galaxy SL. Machine learning, particularly deep learning, can help solve this problem.

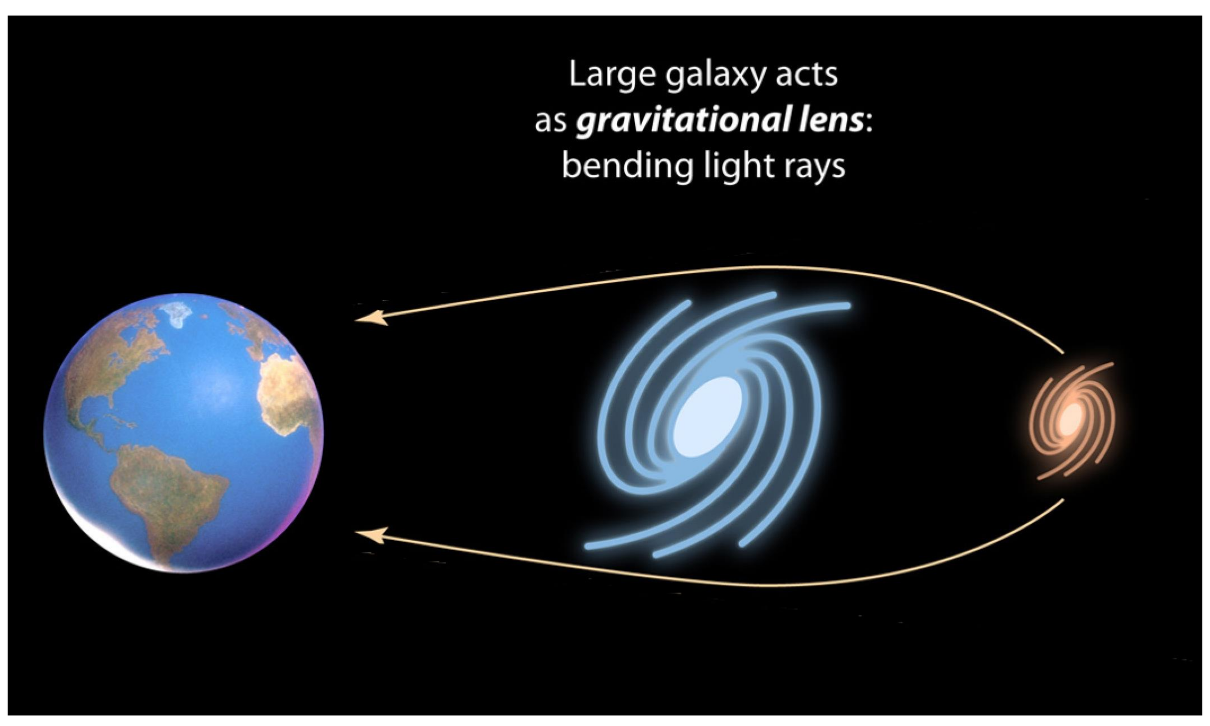

Figure 1. "This diagram illustrates a cosmic phenomenon known as gravitational lensing." Representation of SL due to a galaxy. Credit to Zina Deretsky, NSF (2010)

We want to test and compare many deep learning model architectures with simulated and real data. We intend to build a pipeline to quickly test architectures on benchmark datasets combinations and to identify galaxy-galaxy SL.

Here, we compare the performance of three simulated datasets used for training of one architecture, **Lens Challenge CNN**. A CNN is a deep learning architecture with convolutional and pooling layers for abstraction to learn from visual data.

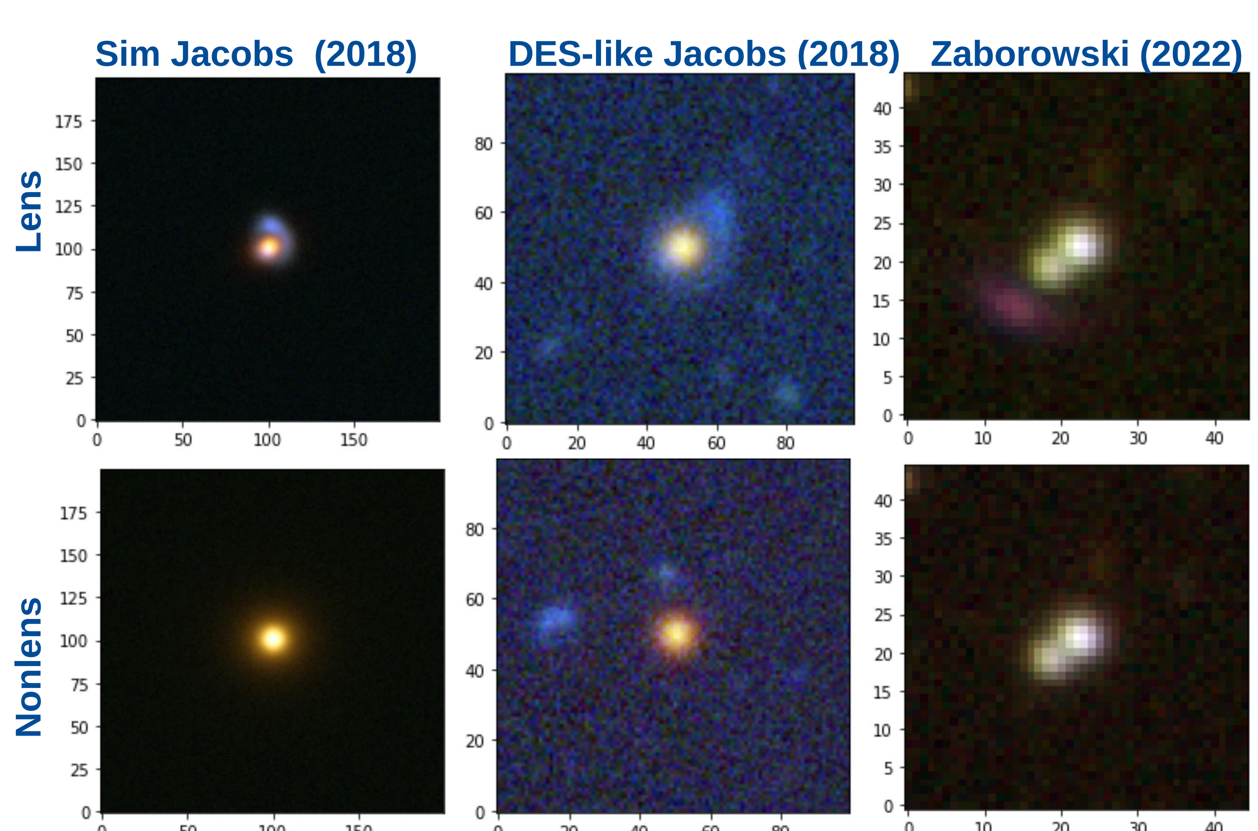

Figure 2. Simulated data examples. First row shows lenses and second row non lenses. Jacobs (2018) data has both without a background nor astronomical sources (first row) and DES-like simulations (second row). Zaborowski (2022) data has simulated lenses over real DELVE images and noise (third row)

Methods and data

We gathered three simulation types (Table 1). Jacobs (2018) has a dataset where both the lens and galaxy are simulated and the background are DES-like simulations, and a dataset with clean simulations without background. Zaborowski (2022) contains simulated lenses over real DELVE imaging.

Collected data was formatted to hdf5 format and plotted as RGB images for visualization purposes (Fig. 2). We then trained the Lens Challenge CNN on all datasets.

Results and future steps

The trained model with best results would have a higher true and lower false positive rate as shown in the Receiver Operating Characteristic (ROC) curve (Fig. 4), and decreasing and close to zero training and validation in the loss curve (Fig. 3).

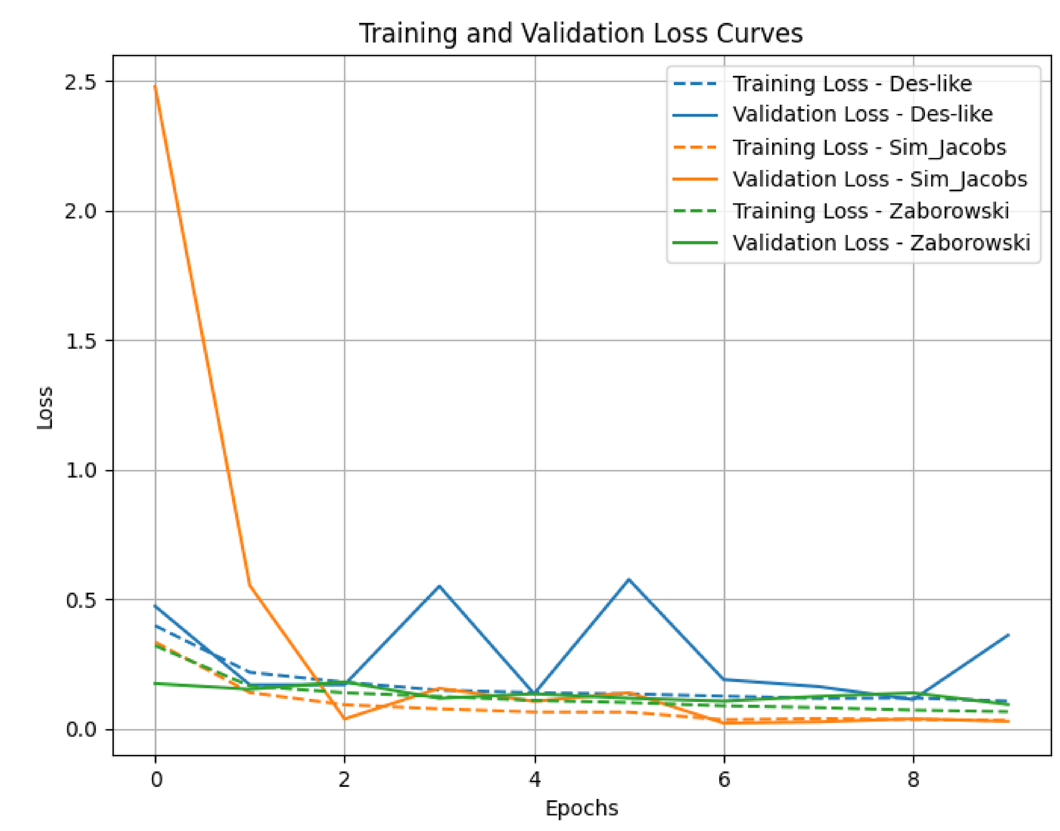

Figure 3. Training and validation loss curve for the three datasets run with the chosen architecture.

	DES-like Jacobs (2018)	Sim Jacobs (2018)	Zaborowski (2022)
Background	Simulated, DES-like	None	Real DELVE images
Total number of images	20,000	20,000	83,286
Image size (px)	100x100	200x200	45x45

Table 1. Information on the simulation data used to run the architecture chosen

Based on this, **Sim Jacobs (2018) gave the best results**, perhaps due to its lack of noise. **Zaborowski (2022) did well** too despite its noisy background and small image size (concerning as images are scaled down and pooled), probably due to being a large dataset. Noisy but with fewer images than Zaborowski, **DES-like Jacobs (2018) understandably did worse**.

In the future, we will account for **domain shifts** — i.e., when training and application data have different probability distributions (Sim Jacobs looks unrealistic, so domain shifts should be considered when used) — and test for the best **data-architecture combinations** for SL detection.

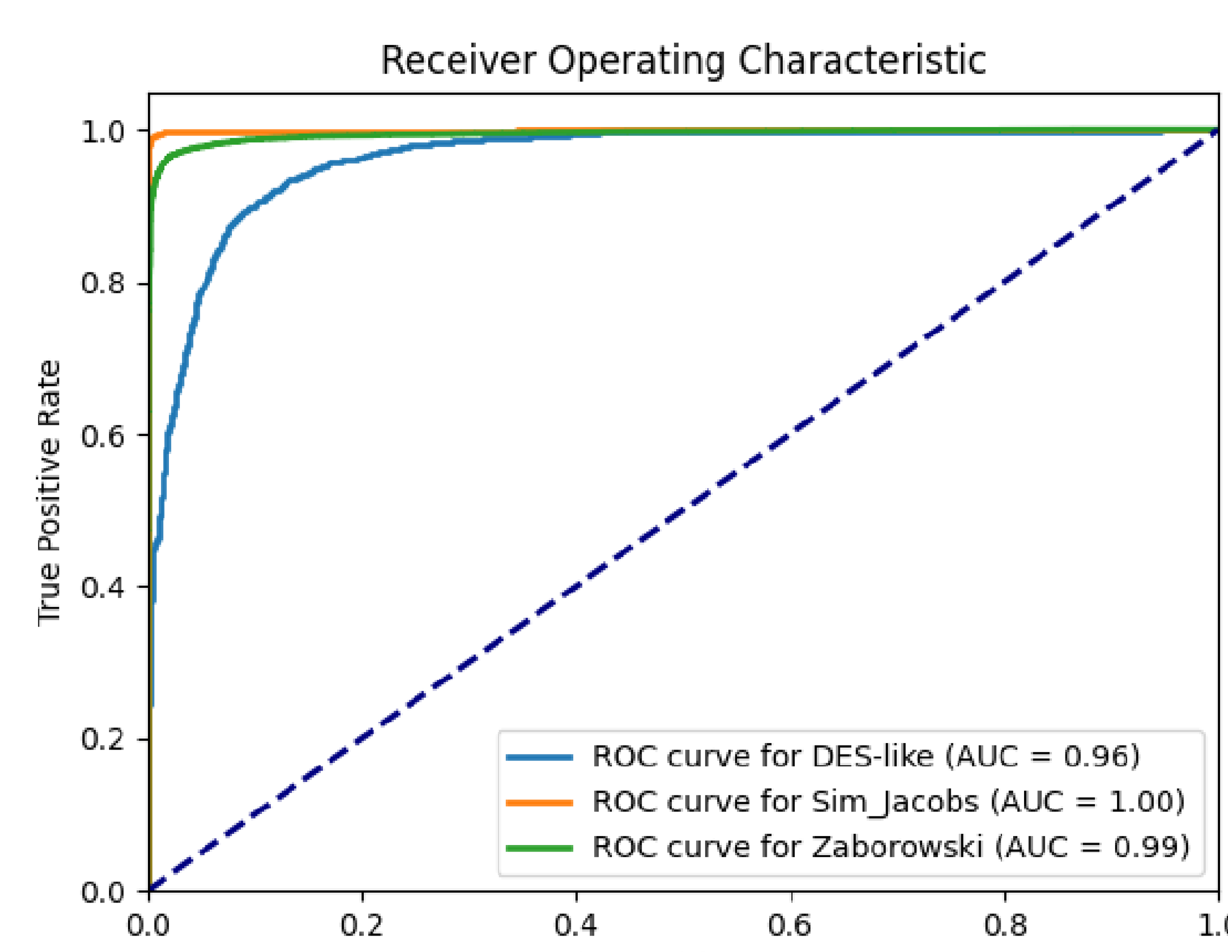

Figure 4. ROC curves, graph with the true positive rate (positives correctly classified as positives) against the false positive rate (negatives incorrectly classified as positives), for the three datasets used to train the chosen Lens Challenge CNN architecture.

References

Jacobs, C. (2018). Finding high-redshift strong lenses in DES using convolutional neural networks. Retrieved from doi.org/10.48550/arXiv.1811.03786
Zaborowski, E. (2022). Identification of Galaxy-Galaxy Strong Lens Candidates in the DECam Local Volume Exploration Survey Using Machine Learning. Retrieved from arxiv.org/abs/2210.10802
Benton, R. (2019). The Strong Gravitational Lens Finding Challenge. Retrieved from arxiv.org/abs/1802.03609

This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

DEEP SKIES