FERMILAB-POSTER-23-227-STUDENT

Benchmarking Neural Network Architectures for Strong Gravitational Lens Classification

Eileen Nolan, George Mason University — SIST Intern | Brian Nord and Alex Drlica-Wagner

Background

- **Strong gravitational lensing** is a phenomenon where a foreground galaxy acts as a powerful lens, bending and magnifying the light emitted from distant background sources.
- **Convolutional Neural Networks** (CNNs) are deep learning models designed for image classification, object detection, and segmentation. CNNs use convolutional layers to detect patterns and features in data.
- Applying CNNs to gravitational lensing **studies requires specialized architectures** and **data handling** due to the scarcity of strong lensing events and data.

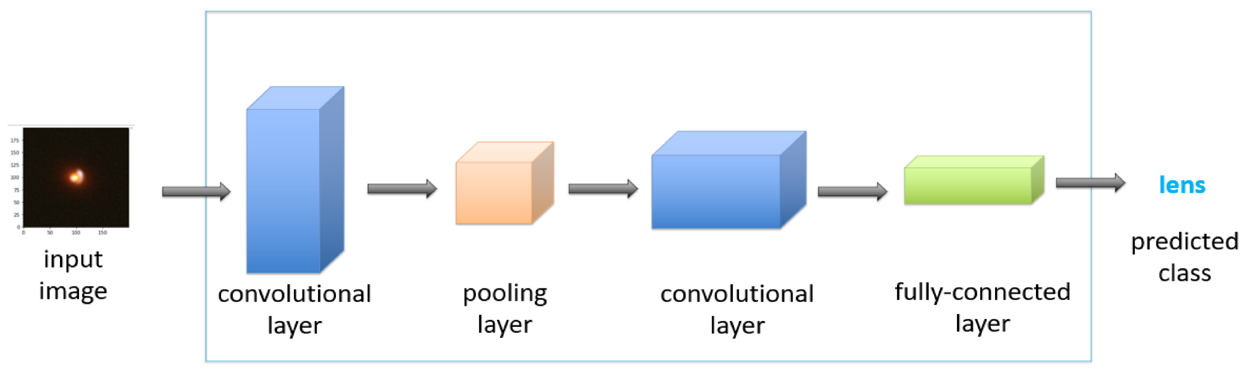


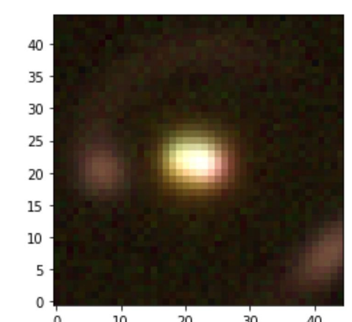
Figure 1: Example of CNN architecture

Motivation & Purpose

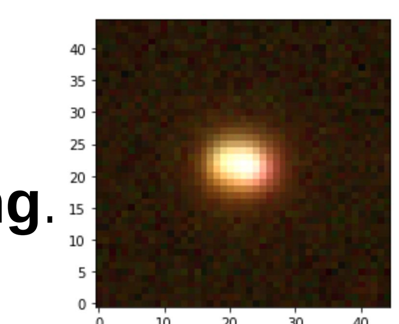
- Detecting more gravitational strong lenses **aids research** on dark energy and dark matter.
- Visual inspection has been a primary method for strong lens detection after automated algorithms were created, indicating our need for improved methods.
- We aim to **explore multiple architectures** and hyperparameters to optimize model performance while minimizing false positives with aim to classify strong lensing.

Data & Methods

- Address a **binary classification problem** by training models on simulated data.
- Test various CNN architectures on one simulated dataset to **compare architecture performance.**
- Gain insight on the optimal architecture features for **classification of strong lensing**.



lens



non-lens

CNN Architectures Notes

- Lens Finder CNN and Galaxy Merger CNN models are **custom CNN implementations** with specific architectures.
- Custom ResNet50 uses a **pre-trained** ResNet-50 model from TensorFlow's Keras API.
- Custom ImageNet is a custom CNN architecture inspired by the ImageNet model.
- All architectures use **ReLU as the activation function** and **Sigmoid as the output activation** function.

References

1. Metcalf, R. B., et al. (2019, March 20). The strong gravitational lens finding challenge. arXiv.org. https://arxiv.org/abs/1802.03609
2. Murray, C., & Cohen, Y. (n.d.). *Hello universe!*. Classifying JWST-HST galaxy mergers with CNNs - Hello Universe! https://spacetelescope.github.io/hellouniverse/notebooks/hello-universe/Classifying_JWST-HST_galaxy_mergers_with_CNNs/Classifying_JWST-HST_galaxy_mergers_with_CNNs.html
3. 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

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.

CNN Architecture Comparison

Table 1: Depicts the similarities and differences between implemented architectures.

	Lens Finder	Galaxy Merger	ResNet50*	ImageNet*
Num Conv Blocks	4	3	5	4
Num Dense	3	2	2	3
Pre-trained	No	No	Yes	No
Dropout	Yes	Yes	No	No
Batch Norm	Yes	Yes	No	Yes

*Indicates those networks have been customized.

Results

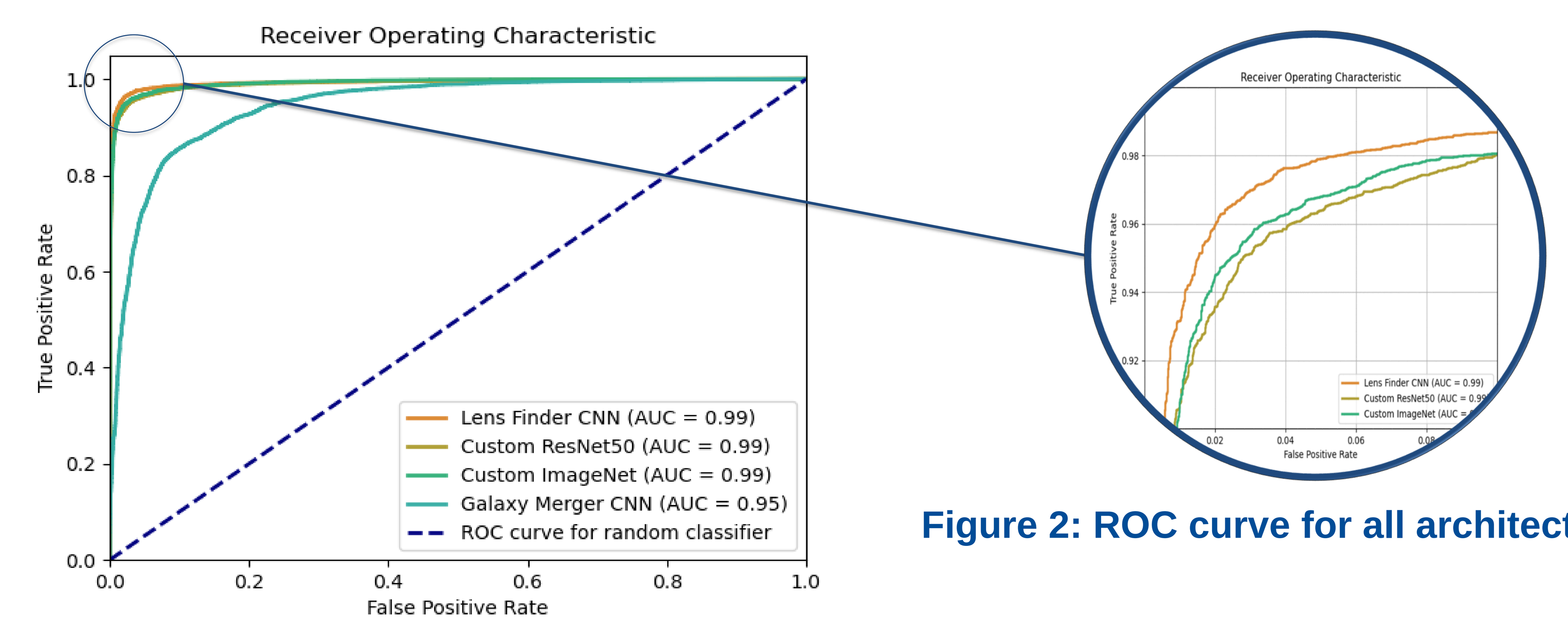


Figure 2: ROC curve for all architectures.

Table 2: Confusion matrix results for all architectures.

	True-Positive	False-Positive	False-Negative	False-Positive
Lens Finder CNN	8445	194	296	7723
Galaxy Merger CNN	8040	559	1545	6474
Custom ResNet50	8383	256	392	7627
Custom ImageNet	8445	194	421	7598

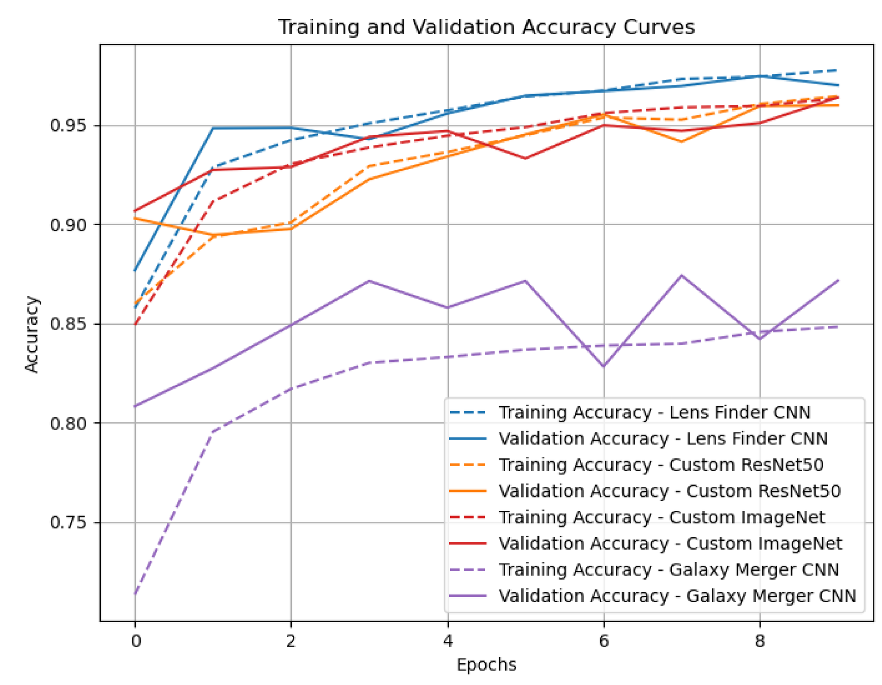


Figure 3: Model Accuracy Curves

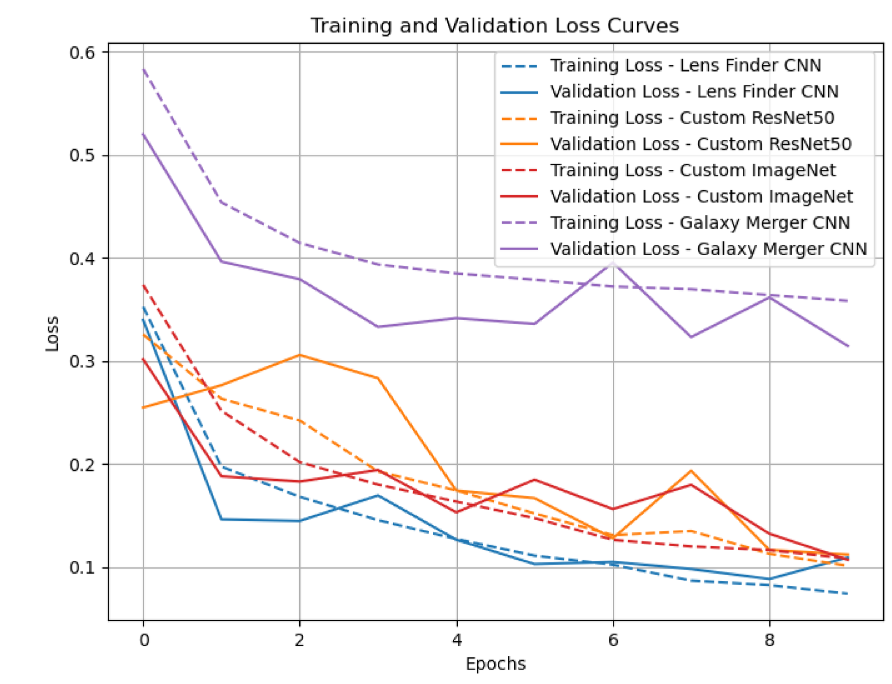


Figure 4: Model Loss Curves

Conclusions

- **Lens Finder CNN** achieved 97% accuracy while minimizing false positive and false negatives. These results suggest the combination of 4 convolutional blocks, 3 deep layers, dropout layers, and batch normalization effectively classify strong lensing in this dataset.
- **Custom ResNet50** and **Custom ImageNet** had similar performance to Lens Finder CNN, indicating that pre-trained models and varying convolutional and dense configurations are also effective for strong lens classification.
- Future work should **individually optimize architectures**' hyperparameters for improved performance and lens classification.
- Future training on **multiple datasets** will enable better comparisons and generalizability.