Searching for Strong Gravitational Lenses in the Dark Energy Survey

Jimena González

OSG All-Hands Meeting
July 10, 2023
OUTLINE

INTRODUCTION
- Strong lensing
- Science applications
- Challenge context

PROJECT WORKFLOW
1 - Training sample
2 - Machine Learning Training
3 - Complete Search

IMPLEMENTING HTC & OSPOOL
- Before attending the OSG school
- Transition & challenges

IMPACT OF HTC & OSG
- Quantitative impact
- Pushing the limits
INTRODUCTION

Strong Lensing Observables:
- Image positions
- Relative magnifications
- Time delays between images

Science Applications:

<table>
<thead>
<tr>
<th>Elliptical Galaxies</th>
<th>Cosmology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass Structure</td>
<td>Dark Energy</td>
</tr>
<tr>
<td>Formation</td>
<td>Dark Matter</td>
</tr>
<tr>
<td>Evolution</td>
<td>Expansion History</td>
</tr>
</tbody>
</table>

Credit: Forbes
Credit: Wikipedia
DES covers 5,000 sq deg. -> 500,000 times larger than the image shown.

Expectation: hundreds of strong lenses in complete data.
1- TRAINING SAMPLE: POSITIVE CLASS

Simulations

We control the quality of the simulations by setting constraints on simulation properties.

Created with Lenstronomy (Birrer & Amara 2018, Birrer et al. 2021)

Computational Limits:

- Need thousands of simulations: ~2 days (for 15,000 simulations)
- Tighter constraints require more time: 2 days -> 4 days
- Couldn’t experiment much with simulations with different properties.
Machine learning model:
- Vision Transformer (ViT): 86 million parameters.

Computational Limits:
- Training is time consuming: $\sim$3 days on personal laptop.
- The larger the training sample, the more time is needed.
  ➡ Each training run is a big commitment.
Search: Must process **320 million** images (1 TB)

**Representation of 72 images**

**Computational Limits:**
- On personal laptop: 7 years
- On a cluster running in series: 4.5 years

➡️ When am I going to graduate?
• Mostly doing simulations, and training. Taking ~2 weeks.
• Applying search to some images (80,000) to understand results. ~2 days
• Spending my time mostly waiting.
• Errors were expensive.

Elephant in the room: search on all the data.
IMPLEMENTING HTC: 1 - SIMULATIONS

How this task is parallelizable:

• Simulations are independent of each other.
• Each simulation only takes ~2 seconds.

Implementing HTC: straight forward

• Organization: input files, output files and software.
• Modify code to receive important information as arguments. Example: number of simulations & random seed.
• Lots of testing!

Expected: 96 h
With HTC: 4 h
MOVING TO OSG: 2 - TRAINING

Challenge: Not easily parallelizable

- The learning process can’t be divided into independent tasks.

OSG resources useful:

- A GPU made a large impact.

Expected: 72 h

With OSG: 4 h
IMPLEMENTING HTC: 3 - TOTAL SEARCH

How this task is parallelizable:

• Images are independent of each other.
• The DES imaging data is divided into ~10,000 pieces (tiles), each containing 24,000 images.
• Processing each tile takes only 4 hours.

Implementing HTC:

• Challenge: All input data is heavy and not public.
• Solution: implementing HTC with the resources I had available: FermiGrid. Same concepts, just different syntax.

Expected: 39420 h
With HTC: 72 h
# Quantitative Impact of the OSG School

## 1- Simulations:

<table>
<thead>
<tr>
<th></th>
<th>Expected</th>
<th>With HTC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>96 h</td>
<td>4 h</td>
</tr>
</tbody>
</table>

## 2- Training:

<table>
<thead>
<tr>
<th></th>
<th>Expected</th>
<th>With OSG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>72 h</td>
<td>4 h</td>
</tr>
</tbody>
</table>

## 3- Complete Search

<table>
<thead>
<tr>
<th></th>
<th>Expected</th>
<th>With HTC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>39420 h</td>
<td>72 h</td>
</tr>
</tbody>
</table>
PUSHING THE LIMITS

Challenge for these searches:
Purity rate of automatic searches: 1%

Possible to implement **Interactive Machine Learning** to improve our training sample:

![Diagram showing the interaction between ML model, human feedback, and new results with 9 categories](image-url)
SEARCH RESULTS

~230 million images → ViT trained model → 20,600 candidates

Random sample of candidates with probability ~1 (~900):

[Images of candidate galaxies]
CONCLUSION

State of the project:
• Preparing a crowdsourcing project in Zooniverse to extract the compelling candidates.
• More plans with OSG: characterizing the final candidates.

Thanks to the OSG consortium for allowing me to focus on the science, instead of the computational limitations.
**EXTRA SLIDES: TRAINING SAMPLE**

**Positive training sample “Single”:**

Ring galaxies
Smooth galaxies
Spiral galaxies
With companions
Crowded fields
Artifacts
Most negatives

**Negative training sample:**

We control the “quality” of the simulations by setting constraints: Magnification, position of the images, contrast of arc simulation, etc.

Created with Lenstronomy (Birrer & Amara 2018, Birrer et al. 2021)
EXTRA SLIDES: THE VISION TRANSFORMER (ViT)

CHARACTERISTICS:
- Larger receptive field
- Better for global features
- Performs the same or better than state of the art CNN models when pre-trained on large datasets.

Class: Bird, ball, car, etc.

Patch + Position Embedding

Linear Projection of Flattened Patches

Transformer Encoder

MLP Head