Intelligence artificielle pour la détection de lentilles gravitationnelles.

Rémi Flamary, Philippa Hartley, Neal Jackson, Amit Tagore, Ben Metcalfe

Journée Scientifique de l’OCA
Sophia Antipolis, November 15 2018
Small history of gravitational lenses

Photon path

\[ \alpha = \frac{4GM}{c^2 \xi} \]

Timeline

1704  Newton suspects gravitational deflection of light.
1915  General relativity predicts twice the deflection of Newton.
1919  Lensing effect observed by Arthur Eddington during a solar eclipse.
1979  Observation of the first strong lens: Twin Quasar Q0957+561A
       [Walsh et al., 1979]
Small history of gravitational lenses

Timeline

1704  Newton suspects gravitational deflection of light.
1915  General relativity predicts twice the deflection of Newton.
1919  Lensing effect observed by Arthur Eddington during a solar eclipse.
1979  Observation of the first strong lens: Twin Quasar Q0957+561A [Walsh et al., 1979]
• Cosmic telescopes (magnification of far-off objects).
• 300 strong lenses currently known, detected by humans.
• Euclid mission [Laureijs et al., 2012], Strong Lens Legacy Science Group: 300,000 galaxy/galaxy lenses out of 10 billion sources. How to find them?
Strong Gravitational lenses

- Cosmic telescopes (magnification of far-off objects).
- 300 strong lenses currently known, detected by humans.
- Euclid mission [Laureijs et al., 2012], Strong Lens Legacy Science Group: 300,000 galaxy/galaxy lenses out of 10 billion sources. How to find them?
Teach the machine to perform a given task.

Give it \( n \) example of observations \( x \) and the corresponding prediction \( y \).

Optimization problem:

\[
\min_f \quad \frac{1}{n} \sum_{i=1}^{n} L(y_i, f(x_i))
\]

We chose Support Vector Machines that work well on small datasets.
Supervised machine learning

- Teach the machine to perform a given task.
- Give it $n$ example of observations $x$ and the corresponding prediction $y$.
- Optimization problem:
  \[
  \min_{f} \frac{1}{n} \sum_{i=1}^{n} L(y_i, f(x_i))
  \]  
  \(1\)
- We chose Support Vector Machines that work well on small datasets.
• Use training data from the lens finding challenge [Metcalf et al., 2018].
• Simulated with Bologna Lens factory.
• 20 k Ground observation (4 wavelengths) and 20k images Space observations.
• Simulated following Kilo degree Survey (Kids) and Euclid observation models.
• Validation on part of the dataset suggest 96% and 88% AUC.
Gravitational lens finding challenge

Competition [Metcalf et al., 2018]

- Training dataset presented earlier.
- 100,000 simulated test images, 48 hours for classifying.
- Performances measured with Area Under the ROC Curve (AUC) and the ratio of correctly classified lenses before a false positive occur (TPR₀).
### Results

- **3 family of submissions:**
  - Convolutional neural networks (CNN).
  - Support vector Machines (us).
  - Human Annotator (us).

- AUC is ability to separate the classes in average.

- CNN works best in AUC, well in TPR$_0$. 

### Table: Competition results in AUC

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>AUROC</th>
<th>TPR$_0$</th>
<th>TPR$_{10}$</th>
<th>Short Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMU-DeepLens-ResNet-ground3</td>
<td>Ground-Based</td>
<td>0.98</td>
<td>0.09</td>
<td>0.45</td>
<td>CNN</td>
</tr>
<tr>
<td>CMU-DeepLens-Resnet-Voting</td>
<td>Ground-Based</td>
<td>0.98</td>
<td>0.02</td>
<td>0.10</td>
<td>CNN</td>
</tr>
<tr>
<td>LASTRO EPFL</td>
<td>Ground-Based</td>
<td>0.97</td>
<td>0.07</td>
<td>0.11</td>
<td>CNN</td>
</tr>
<tr>
<td>CAS Swinburne Melb</td>
<td>Ground-Based</td>
<td>0.96</td>
<td>0.02</td>
<td>0.08</td>
<td>CNN</td>
</tr>
<tr>
<td>AstrOmatic</td>
<td>Ground-Based</td>
<td>0.96</td>
<td>0.00</td>
<td>0.01</td>
<td>CNN</td>
</tr>
<tr>
<td>Manchester SVM</td>
<td>Ground-Based</td>
<td>0.93</td>
<td>0.22</td>
<td>0.35</td>
<td>SVM / Gabor</td>
</tr>
<tr>
<td>Manchester-NA2</td>
<td>Ground-Based</td>
<td>0.89</td>
<td>0.00</td>
<td>0.01</td>
<td>Human Inspection</td>
</tr>
<tr>
<td>ALL-star</td>
<td>Ground-Based</td>
<td>0.84</td>
<td>0.01</td>
<td>0.02</td>
<td>edges/gradients and Logistic Reg.</td>
</tr>
<tr>
<td>CAST</td>
<td>Ground-Based</td>
<td>0.83</td>
<td>0.00</td>
<td>0.00</td>
<td>CNN / SVM</td>
</tr>
<tr>
<td>YattaLensLite</td>
<td>Ground-Based</td>
<td>0.82</td>
<td>0.00</td>
<td>0.00</td>
<td>SExtractor</td>
</tr>
</tbody>
</table>
Results

- 3 family of submissions:
  - Convolutional neural networks (CNN).
  - Support vector Machines (us).
  - Human Annotator (us).
- AUC is ability to separate the classes in average.
- CNN works best in AUC, well in TPR₀.
Competition results in $\text{TPR}_0$

<table>
<thead>
<tr>
<th>Name</th>
<th>type</th>
<th>AUROC</th>
<th>$\text{TPR}_0$</th>
<th>$\text{TPR}_{10}$</th>
<th>short description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manchester SVM</td>
<td>Ground-Based</td>
<td>0.93</td>
<td>0.22</td>
<td>0.35</td>
<td>SVM / Gabor</td>
</tr>
<tr>
<td>CMU-DeepLens-ResNet-ground3</td>
<td>Ground-Based</td>
<td>0.98</td>
<td>0.09</td>
<td>0.45</td>
<td>CNN</td>
</tr>
<tr>
<td>LASTRO EPFL</td>
<td>Ground-Based</td>
<td>0.97</td>
<td>0.07</td>
<td>0.11</td>
<td>CNN</td>
</tr>
<tr>
<td>CMU-DeepLens-Resnet-Voting</td>
<td>Ground-Based</td>
<td>0.98</td>
<td>0.02</td>
<td>0.10</td>
<td>CNN</td>
</tr>
<tr>
<td>CAS Swinburne Melb</td>
<td>Ground-Based</td>
<td>0.96</td>
<td>0.02</td>
<td>0.08</td>
<td>CNN</td>
</tr>
<tr>
<td>ALL-star</td>
<td>Ground-Based</td>
<td>0.84</td>
<td>0.01</td>
<td>0.02</td>
<td>edges/graditants and Logistic Reg.</td>
</tr>
<tr>
<td>Manchester-NA2</td>
<td>Ground-Based</td>
<td>0.89</td>
<td>0.00</td>
<td>0.01</td>
<td>Human Inspection</td>
</tr>
<tr>
<td>YattaLensLite</td>
<td>Ground-Based</td>
<td>0.82</td>
<td>0.00</td>
<td>0.00</td>
<td>SExtractor</td>
</tr>
<tr>
<td>CAST</td>
<td>Ground-Based</td>
<td>0.83</td>
<td>0.00</td>
<td>0.00</td>
<td>CNN / SVM</td>
</tr>
<tr>
<td>AstrOmatic</td>
<td>Ground-Based</td>
<td>0.96</td>
<td>0.00</td>
<td>0.01</td>
<td>CNN</td>
</tr>
</tbody>
</table>

Results

- $\text{TPR}_0$ is the ratio of detected lenses before a false positive occur when sorted by classifier scores.
- Measure of trust for the highest predicted scores, better for retrieval.
- SVM work far better in $\text{TPR}_0$ for space data.
- CNN work better on space data.
- None of the methods is designed to optimize this criterion.
Man vs machine

Results

- Eyeball inspection of the 100,000 simulated test images.
- 5 level confidence score.
- Done by Neal Jackson and Amit Tagore (5000/2000 imgs/h).
Results

- Eyeball inspection of the 100,000 simulated test images.
- 5 level confidence score.
- Done by Neal Jackson and Amit Tagore (5000/2000 imgs/h).
- Greatly outmatched by the best automatic methods (CNN/SVM).
- Human and machine have different strength (small/large Einstein radius).
Man vs machine

Results

- Eyeball inspection of the 100,000 simulated test images.
- 5 level confidence score.
- Done by Neal Jackson and Amit Tagore (5000/2000 imgs/h).
- Greatly outmatched by the best automatic methods (CNN/SVM).
- Human and machine have different strength (small/large Einstein radius).
Results

- Eyeball inspection of the 100 000 simulated test images.
- 5 level confidence score.
- Done by Neal jackson and Amit Tagore (5000/2000 imgs/h).
- Greatly outmatched by the best automatic methods (CNN/SVM).
- Human and machine have different strength (small/large Einstein radius).
- Apply our classifier on 1 million images from Kilo Degree Survey (KiDS).
- Classification score far more uncertain (simulation $\neq$ real life).
- Look at images with larger scores.
- Kept 213 object that range from possibly to very likely lenses.
Apply our classifier on 1 million images from Kilo Degree Survey (KiDS).
Classification score far more uncertain (simulation $\neq$ real life).
Look at images with larger scores.
Kept 213 object that range from possibly to very likely lenses.
Conclusions

• We need automatic procedure to detect strong gravitational lenses.
• Machines now surpass humans in finding lenses.
• Strength of SVMs when false positives are a problem.
• CNN better approach (they learn the Gabor filters).

Best strategy?

• Use CNN but encode expert knowledge (polar representation, ...)
• Design dedicated objective to minimize false positives (neyman-pearson classification)
• Discrepancy between training and test data?
Domain adaptation (special case of transfer learning)

- Problem: New data is different from training data.
- In astronomy: Simulated data is always different from real life data.
- **How to train on simulated data but still work on real data?**
- Use of Optimal Transport theory to adapt between domains [Courty et al., 2016, Damodaran et al., 2018].
Domain adaptation (special case of transfer learning)

- Problem: New data is different from training data.
- In astronomy: Simulated data is always different from real life data.
- How to train on simulated data but still work on real data?
- Use of Optimal Transport theory to adapt between domains [Courty et al., 2016, Damodaran et al., 2018].
Thank you
**Optimal transport for domain adaptation.**
*Pattern Analysis and Machine Intelligence, IEEE Transactions on.*

**Deepjdot: Deep joint distribution optimal transport for unsupervised domain adaptation.**
In *European Conference in Computer Visions (ECCV).*

Euclid: Esa’s mission to map the geometry of the dark universe.

In *Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave*, volume 8442, page 84420T. International Society for Optics and Photonics.


The strong gravitational lens finding challenge.

Technical report.

0957 + 561 A, B - Twin quasistellar objects or gravitational lens.  