Hunting for Rare Treasure with Machine Learning: Gravitational Lensing and Glitches in the Euclid space telescope and Vera Rubin Observatory

Supervision team: Prof Stephen Serjeant, Dr Hugh Dickinson

External supervisor: Dr Jane Bromley

Lead contact: Prof Stephen Serjeant

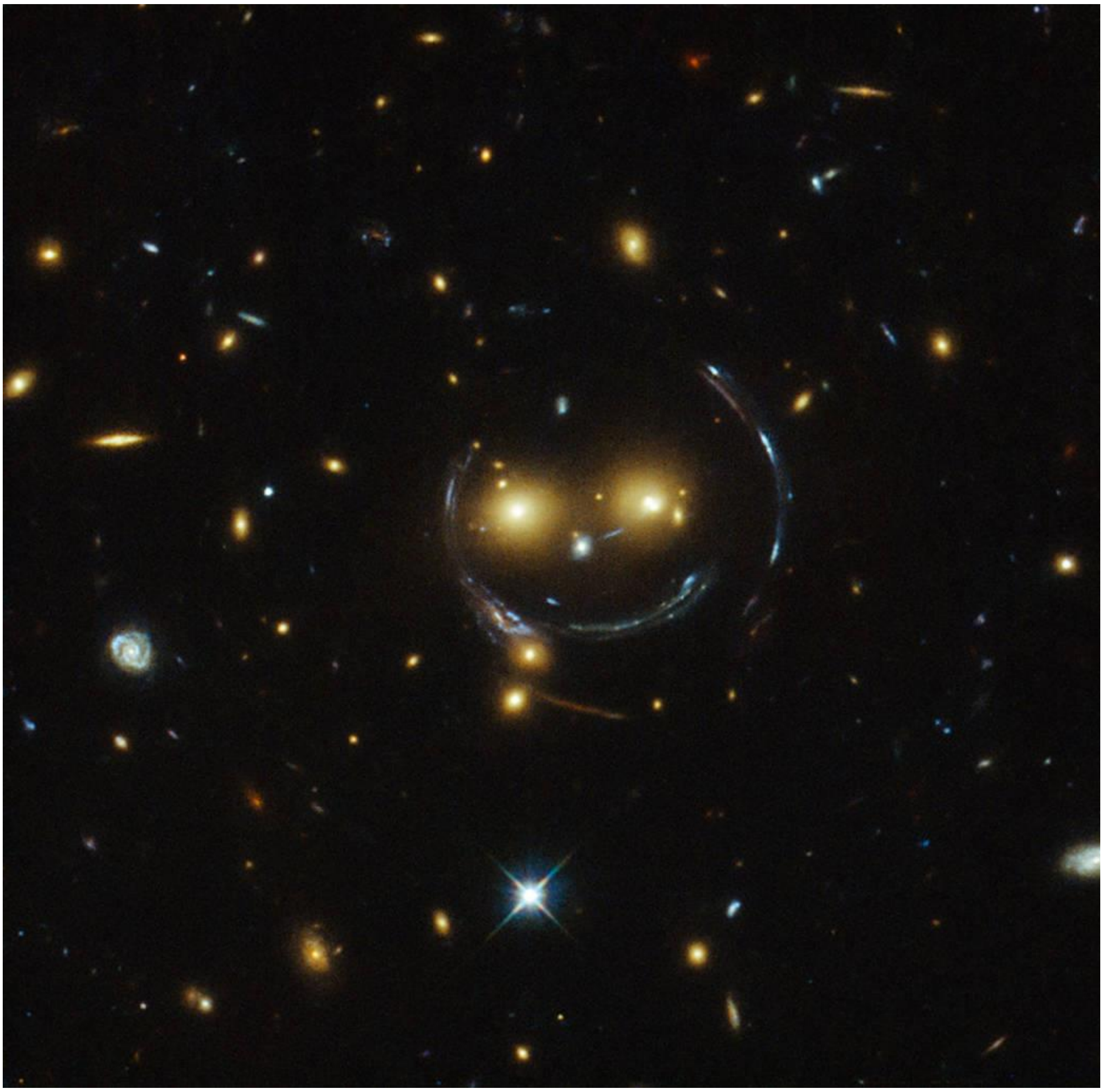

Image Credit: NASA/ESA
Image source: https://www.nasa.gov/content/hubble-sees-a-smiling-lens

Project highlights:

- Project importance: to find large, pure samples of rare strong gravitational lenses in forthcoming giant astronomical surveys, and in particular rare "jackpot" lenses from

which dark matter halo profiles can be reconstructed and independent geometrical constraints can be made of the dark matter equation of state.
- Outcomes: self-supervised machine learning to find rare events in future large astronomical surveys
- Career development, learning and training opportunities: DISCnet studentships provide students with a very wide range of training and networking opportunities, including: residential training courses on machine learning, handling large data, high performance computing, statistics; two funded 3-month industrial placements to deploy the techniques in the PhD in an industrial context, in addition to the 3.5 year PhD duration; career development networking events with industry and academia.
- Partners and collaborators: the project will involve joining working groups of the Euclid space telescope (launching in 2022) and the LSST project on the Vera Rubin Observatory, which will give the student visibility in large international research consortia to promote their professional development.

Project description:
Strong gravitational lensing is where there is a chance alignment of a foreground and a background galaxy, so the background galaxy is viewed through the warped spacetime of the foreground one. The background galaxy can be magnified by factors of a few to over a hundred, making strong lensing a powerful window on distant, faint galaxy evolution, as well as tracking the dark matter halo profiles around the foreground galaxies. There are of the order 1000 strong lensing systems discovered to date, but forthcoming extragalactic surveys with the Euclid space telescope and the Vera Rubin Observatory are poised to revolutionise the study of strong gravitational lenses, with of the order 100,000 strong lenses waiting to be discovered. However, these 100,000 lensing systems will be interspersed among around a billion non-lensed galaxies, so this is a non-trivial data mining problem. In general, a major advantage of large catalogues in astronomy is the ability to find rare objects, which in this case includes "jackpot" lenses, also known as compound lenses, where several galaxies at different redshifts are lined up along the line of sight. These jackpot lenses are very useful for constraining dark matter halo profiles, a key test of cosmological models, and they can provide a measure of the dark energy equation of state (arguably the most pressing fundamental physics measurement needed today), but finding a rare subset among the already-rare strong lenses is a "needle in a haystack in a field full of haystacks" problem.

This has led us, and several other groups, to develop machine learning technologies for parsing the large data sets. At the OU we have trained convolutional neural nets for these purposes and developed techniques for interpreting the workings of these ostensible "black box" algorithms. We would like to recruit a PhD student to continue this work, to using self-supervised learning using similarity metrics based on the outputs at the bottleneck of an autoencoder. This would help solve the problem of having small training sets for particular rare lensing configurations, such as jackpot lenses or ultra-high-redshift lensing systems. We also would like to use established techniques to help identify where in an image the strong lensing signal is coming from. A similar rare object detection technique, but based on non-similarity, can also be used to identify non-astrophysical glitches in early commissioning data from Euclid and the Vera Rubin Observatory. There are many directions this project can go in, both in the development of machine learning technologies and their applications to major new data sets, as well as in astrophysical follow-ups of the strong lensing systems discovered in the course of the project.

References:

1. Wilde et al. 2021 MNRAS submitted (arXiv link to be added)
2. Davies et al. 2019 MNRAS 487, 5263 https://ui.adsabs.harvard.edu/abs/2019MNRAS.487.5263D/abstract
3. Lauritsen et al. 2021 MNRAS 507, 1546 https://ui.adsabs.harvard.edu/abs/2021MNRAS.507.1546L/abstract

Qualifications required:

Minimum requirements as per DISCnet studentships:

At least BSc 2:1 or an MSci or MPhys in physics, astrophysics or a relevant discipline