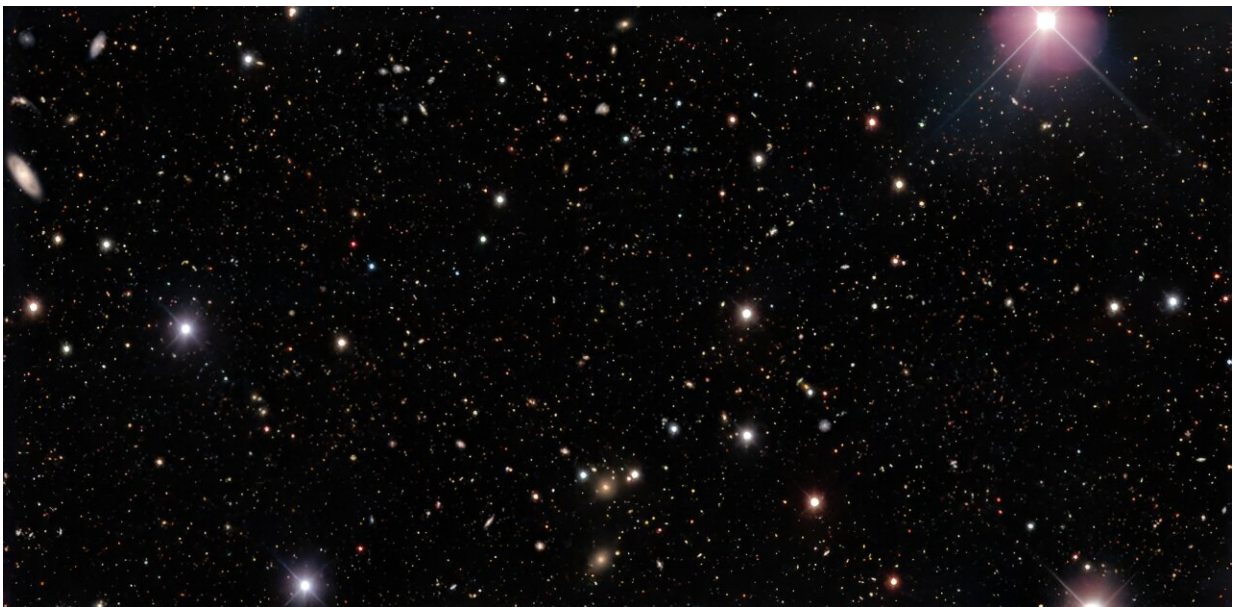


Astronomers apply machine learning techniques to find early-universe quasars in an ocean of data

July 11 2024, by Trisha Muro



A deep image from the Dark Energy Survey showing the field covered by one of the individual detectors in the Dark Energy Camera. Credit: DES Collaboration/NOIRLab/NSF/AURA/M. Zamani

Quasars are extremely luminous galactic cores where gas and dust falling into a central supermassive black hole emit enormous amounts of light. Due to their exceptional brightness, these objects can be seen at high redshifts, i.e., large distances.

A higher redshift not only indicates a quasar is at a greater distance but also further back in time. Astronomers are interested in studying these ancient objects because they hold clues about the evolution of our universe in its early adolescence.

High-redshift quasar candidates are initially identified by their color—they are very red—and must then be confirmed as such by looking at separate observations of their spectra. However, some high-redshift candidates can be mistakenly eliminated from further investigation because of distortions in their appearance caused by gravitational lensing.

This is a phenomenon that occurs when a [massive object](#), such as a galaxy, is located between us and a distant object. The galaxy's mass bends space to act a bit like a magnifying glass, causing the path taken by the distant object's light to be bent and resulting in a distorted image of the object.

While this alignment can be beneficial—the gravitational lens magnifies the image of the quasar, making it brighter and easier to detect—it can also deceptively alter the quasar's appearance.

Interfering light from the stars in the intervening lensing galaxy can make the quasar appear more blue, while the bending of spacetime can make it appear smeared or multiplied. Both of these effects make it likely to be eliminated as a quasar candidate.

So a team of astronomers led by Xander Byrne, astronomer at the University of Cambridge and lead author of the [paper](#) presenting these results in the *Monthly Notices of the Royal Astronomical Society*, set out to recover the lensed quasars that were overlooked in previous surveys.

Byrne went hunting for these missing treasures in the extensive data

archive from the Dark Energy Survey (DES). DES was conducted with the Department of Energy-fabricated Dark Energy Camera, mounted on the Víctor M. Blanco 4-meter Telescope at the U.S. National Science Foundation Cerro Tololo Inter-American Observatory, a Program of NSF NOIRLab.

The challenge, then, was to devise a way to uncover these cosmic gems from within the enormous ocean of data.

The full DES dataset includes more than 700 million objects. Byrne pared down this archive by comparing the data with images from other surveys to filter out unlikely candidates, including objects that were likely brown dwarfs, which, despite being utterly different from quasars in almost every way, can look surprisingly similar to quasars in images. This process yielded a much more manageable dataset containing 7,438 objects.

Byrne needed to maximize efficiency as he searched those 7,438 objects, but he knew that traditional techniques would likely miss the high-redshift lensed quasars he sought. "To avoid casting out lensed quasars prematurely we applied a contrastive learning algorithm and it worked like a charm."

Contrastive learning is a type of artificial intelligence (AI) algorithm in which sequential decisions place each data point into a group according to what it is or what it is not. "It may seem like magic," said Byrne, "but the algorithm uses no more information than what is already there in the data. Machine learning is all about finding which bits of data are useful."

Byrne's decision to not rely on human visual interpretation led him to consider an unsupervised AI process, meaning the algorithm itself drives the learning process rather than a human.

Supervised [machine learning](#) algorithms are based on a so-called ground-level truth, defined by a human programmer. For example, the process might start with a description of a cat and move through decisions such as "This is/is not an image of a cat. This is/is not an image of a black cat."

In contrast, unsupervised algorithms do not rely on that initial, human-specified definition as the basis for its decisions. Instead, the algorithm sorts each data point according to similarities to the other data points in the set. Here, the algorithm would find similarities among images of multiple animals and would group them as cat, dog, giraffe, penguin, etc.

Beginning with Byrne's 7,438 objects, the unsupervised algorithm sorted the objects into groups. Embracing a geographical analogy, the team referred to the groupings of data as an archipelago. (The term does not imply any proximity in space between objects. It is their characteristics that group them "close" together, not their positions in the sky.)

Within this archipelago, a small "island" subset of objects were grouped together as possible quasar candidates. Among those candidates, four stood out like gems in a pile of pebbles.

Using archival data from the Gemini South telescope, one half of the International Gemini Observatory, which is operated by NSF NOIRLab, Byrne confirmed that 3 of the 4 candidates on "quasar island" are indeed high-redshift [quasars](#). And one of those is very likely to be the cosmic bounty that Byrne was hoping to find—a gravitationally lensed high-redshift quasar. The team is now planning follow-up imaging to confirm the lensed nature of the quasar.

"If confirmed, the discovery of one lensed quasar in a sample of four targets would be a remarkably high success rate! And if this search had been conducted using standard search methods, it's likely this gem would

have remained hidden."

Byrne's work serves as a clever example of how AI might aid astronomers as they search through increasingly larger treasure chests of data. Massive influxes of astronomical data are expected in the coming years with the Dark Energy Spectroscopic Instrument's ongoing five-year survey, as well as the upcoming Legacy Survey and Space and Time, which will be conducted by Vera C. Rubin Observatory beginning in 2025.

More information: Xander Byrne et al, Quasar Island – three new $z \sim 6$ quasars, including a lensed candidate, identified with contrastive learning, *Monthly Notices of the Royal Astronomical Society* (2024). [DOI: 10.1093/mnras/stae902](https://doi.org/10.1093/mnras/stae902)

Provided by NSF's NOIRLab

Citation: Astronomers apply machine learning techniques to find early-universe quasars in an ocean of data (2024, July 11) retrieved 11 July 2024 from <https://phys.org/news/2024-07-astronomers-machine-techniques-early-universe.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--