27 million galaxy morphologies quantified and cataloged with the help of machine learning

April 12 2021, by Erica K. Brockmeier

An image of NGC 1365 collected by the Dark Energy Survey. Also known as the Great Barred Spiral Galaxy, NGC 1365 is an example of a spiral galaxy and is located about 56 million light-years away. Credit: DECam, DES Collaboration

Research from Penn's Department of Physics and Astronomy has produced the largest catalog of galaxy morphology classification to date. Led by former postdocs Jesús Vega-Ferrero and Helena Domínguez Sánchez, who worked with professor Mariangela Bernardi, this catalog of 27 million galaxy morphologies provides key insights into the evolution of the universe. The study was published in *Monthly Notices of the Royal Astronomical Society*.

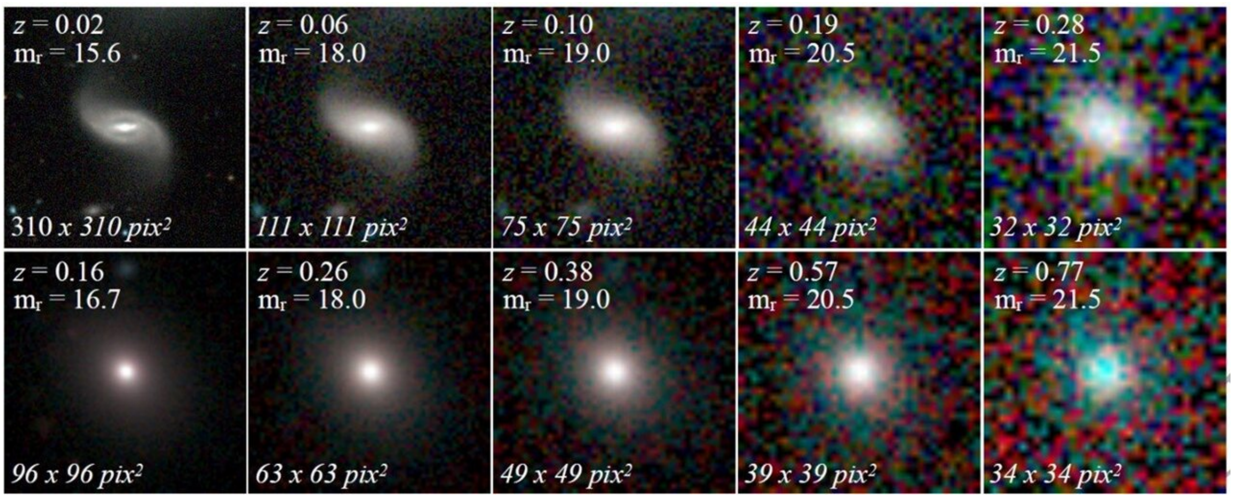
The researchers used data from the Dark Energy Survey (DES), an international research program whose goal is to image one-eighth of the sky to better understand dark energy's role in the accelerating expansion of the universe.

A byproduct of this survey is that the DES data contains many more images of distant galaxies than other surveys to date. "The DES images show us what galaxies looked like more than 6 billion years ago," says Bernardi.

And because DES has millions of high-quality images of astronomical objects, it's the perfect dataset for studying galaxy morphology. "Galaxy morphology is one of the key aspects of galaxy evolution. The shape and structure of galaxies has a lot of information about the way they were formed, and knowing their morphologies gives us clues as to the likely pathways for the formation of the galaxies," Domínguez Sánchez says.

Previously, the researchers had published a morphological catalog for more than 600,000 galaxies from the Sloan Digital Sky Survey (SDSS). To do this, they developed a convolutional neural network, a type of machine learning algorithm, that was able to automatically categorize whether a galaxy belonged to one of two major groups: spiral galaxies, which have a rotating disk where new stars are born, and elliptical galaxies, which are larger, and made of older stars which move more randomly than their spiral counterparts.

But the [catalog](#) developed using the SDSS dataset was primarily made of bright, nearby galaxies, says Vega-Ferrero. In their latest study, the researchers wanted to refine their [neural network model](#) to be able to classify fainter, more distant galaxies. "We wanted to push the limits of morphological classification and trying to go beyond, to fainter objects or objects that are farther away," Vega-Ferrero says.



Images of a simulated spiral (top) and elliptical galaxy at varying image quality and redshift levels, illustrating how fainter and more distant galaxies might look within the DES dataset. Credit: Jesus Vega-Ferrero and Helena Dominguez-Sanchez

To do this, the researchers first had to train their neural network model to be able to classify the more pixelated images from the DES dataset. They first created a training model with previously known morphological classifications, comprised of a set of 20,000 galaxies that overlapped between DES and SDSS. Then, they created simulated versions of new galaxies, mimicking what the images would look like if they were farther away using code developed by staff scientist Mike Jarvis.

Once the model was trained and validated on both simulated and real galaxies, it was applied to the DES dataset, and the resulting catalog of 27 million galaxies includes information on the probability of an individual galaxy being elliptical or spiral. The researchers also found that their neural network was 97% accurate at classifying galaxy morphology, even for galaxies that were too faint to classify by eye.

"We pushed the limits by three orders of magnitude, to objects that are 1,000 times fainter than the original ones," Vega-Ferrero says. "That is why we were able to include so many more galaxies in the catalog."

"Catalogs like this are important for studying galaxy formation," Bernardi says about the significance of this latest publication. "This catalog will also be useful to see if the morphology and stellar populations tell similar stories about how galaxies formed."

For the latter point, Domínguez Sánchez is currently combining their morphological estimates with measures of the chemical composition, age, star-formation rate, mass, and distance of the same galaxies. Incorporating this information will allow the researchers to better study the relationship between galaxy morphology and star formation, work that will be crucial for a deeper understanding of galaxy evolution.

Bernardi says that there are a number of open questions about galaxy evolution that both this new catalog, and the methods developed to create it, can help address. The upcoming LSST/Rubin survey, for example, will use similar photometry methods to DES but will have the capability of imaging even more distant objects, providing an opportunity to gain even deeper understanding of the evolution of the universe.

More information: J Vega-Ferrero et al. Pushing automated morphological classifications to their limits with the Dark Energy

Survey, *Monthly Notices of the Royal Astronomical Society* (2021). DOI: 10.1093/mnras/stab594

Provided by University of Pennsylvania