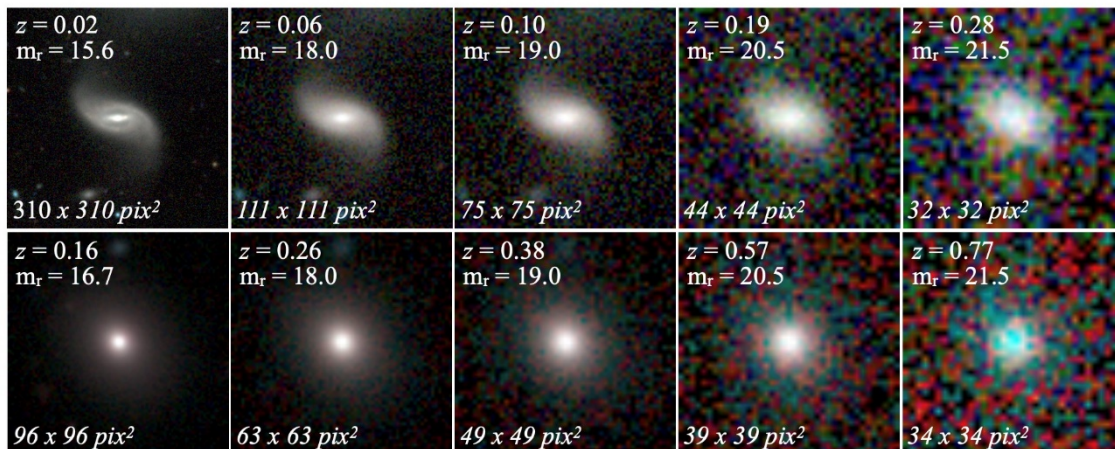




Madrid, lunes 22 de marzo de 2021

## El CSIC participa en el mayor catálogo de clasificación morfológica de galaxias hasta la fecha

- Este catálogo incluye 27 millones de galaxias
- El estudio emplea un algoritmo de aprendizaje automático con hasta un 97% de precisión para clasificar las galaxias, incluso las más débiles y lejanas



Imágenes originales (paneles a la izquierda) de una galaxia espiral (arriba) y una galaxia elíptica (abajo) y sus versiones degradadas, utilizadas para entrenar la red neuronal. / J. Vega-Ferrero *et al.*

Una investigación en la que ha participado el Consejo Superior de Investigaciones Científicas (CSIC) ha producido el mayor catálogo de clasificación morfológica de galaxias hasta la fecha, que incluye **27 millones de galaxias**. Este catálogo, que combina imágenes de alta calidad con mecanismos de inteligencia artificial y en el que han participado investigadores del Instituto de Física de Cantabria (IFCA-CSIC-UC) y el Instituto de Ciencias del Espacio (ICE-CSIC), ha sido publicado en la revista [Monthly Notices of the Royal Astronomical Society \(MNRAS\)](#).

Los investigadores han utilizado los datos de la Exploración de Energía Oscura (Dark Energy Survey, DES) –que catalogó cientos de millones de galaxias distantes durante seis años– y un **algoritmo de aprendizaje automático con hasta un 97% de precisión** para

aprender a clasificar las galaxias en dos tipos de morfologías, incluso las galaxias más débiles y lejanas.

La morfología de las galaxias está muy relacionada con el tipo de estrellas que las componen y sus mecanismos de formación. Principalmente, este catálogo clasifica las galaxias en dos tipos de morfologías: galaxias espirales, que tienen un disco giratorio donde nacen nuevas estrellas, y elípticas, que son las galaxias más masivas del Universo, compuestas de estrellas antiguas que realizan movimientos aleatorios.

Aunque resulta fácil distinguir estos dos tipos de galaxias a simple vista, existen dos problemas importantes: por un lado, el elevado número de galaxias por clasificar, que hizo necesario realizar **clasificaciones automatizadas**, y, por otro lado, el hecho de que las galaxias situadas a mayor distancia parecen más débiles y más pequeñas, así que las imágenes recabadas solían tener mucho ruido.

El equipo científico ha degradado las imágenes de alta calidad de las galaxias locales hasta obtener la apariencia que tendrían si estuvieran más distantes y han introducido las etiquetas correctas para entrenar una **red neuronal convolucional**. De esta manera, ha sido posible aprender a clasificar incluso los ejemplos más difíciles. El estudio señala que **el algoritmo utilizado puede acertar la morfología de las galaxias hasta el 97% de las veces**, independientemente del nivel de ruido y la resolución espacial de las imágenes.

“Este trabajo demuestra que las máquinas son capaces de recuperar imágenes que el ojo humano no puede captar y que tienen capacidad para distinguir las señales útiles del ruido cuando se entrenan con las etiquetas correctas. Por lo tanto, pueden clasificar de forma fiable imágenes de galaxias más débiles”, explica la investigadora Helena Domínguez, del Instituto de Ciencias del Espacio.

El uso de redes neuronales convolucionales es extremadamente exitoso para analizar y clasificar imágenes de galaxias. Este tipo de redes neuronales son un algoritmo de aprendizaje automático (*deep learning*) que puede recibir una imagen de entrada y asignarles una etiqueta a distintos aspectos de dicha imagen y diferenciarlos entre sí.

## Implicaciones futuras

Con este método automatizado, ha sido posible asignar una clasificación a 27 millones de galaxias y elaborar el mayor catálogo morfológico de galaxias publicado hasta la fecha.

Algunas de las galaxias incluidas en el catálogo se encuentran a una distancia de hasta 8 giga-años (Ga), es decir, 8.000 millones de años. Este catálogo permite tener una **visión aproximada de cómo eran las galaxias cuando el Universo tenía la mitad de su edad actual**, estudiar cómo han cambiado las galaxias en los últimos 8 Ga y observar cómo esos cambios estructurales tienen que ver con los caminos evolutivos de las galaxias.

El hecho de que las máquinas puedan aprender a reconocer patrones en datos con ruido y difíciles de interpretar puede tener **aplicaciones directas en otros campos**, como la **seguridad** (por ejemplo, reconocimiento facial), la **industria del reconocimiento de imágenes**, **diagnósticos clínicos** o el **cambio climático**.

J Vega-Ferrero, H Domínguez Sánchez, M Bernardi, M Huertas-Company, R Morgan, B Margalef, M Agüena, S Allam, J Annis, S Avila, D Bacon, E Bertin, D Brooks, A Carnero Rosell, M Carrasco Kind, J Carretero, A Choi, C Conselice, M Costanzi, L N da Costa, M E S Pereira, J De Vicente, S Desai, I Ferrero, P Fosalba, J Frieman, J García-Bellido, D Gruen, R A Gruendl, J Gschwend, G Gutierrez, W G Hartley, S R Hinton, D L Hollowood, K Honscheid, B Hoyle, M Jarvis, A G Kim, K Kuehn, N Kuropatkin, M Lima, M A G Maia, F Menanteau, R Miquel, R L C Ogando, A Palmese, F Paz-Chinchón, A A Plazas, A K Romer, E Sanchez, V Scarpine, M Schubnell, S Serrano, I Sevilla-Noarbe, M Smith, E Suchyta, M E C Swanson, G Tarle, F Tarsitano, C To, D L Tucker, T N Varga, R D Wilkinson. **Pushing automated morphological classifications to their limits with the Dark Energy Survey.** *Monthly Notices of the Royal Astronomical Society.*

Comunicación ICE / CSIC Comunicación