Detecting Anomalous Galaxies with Generative Adversarial Networks

Kate Storey-Fisher, Department of Physics, New York University

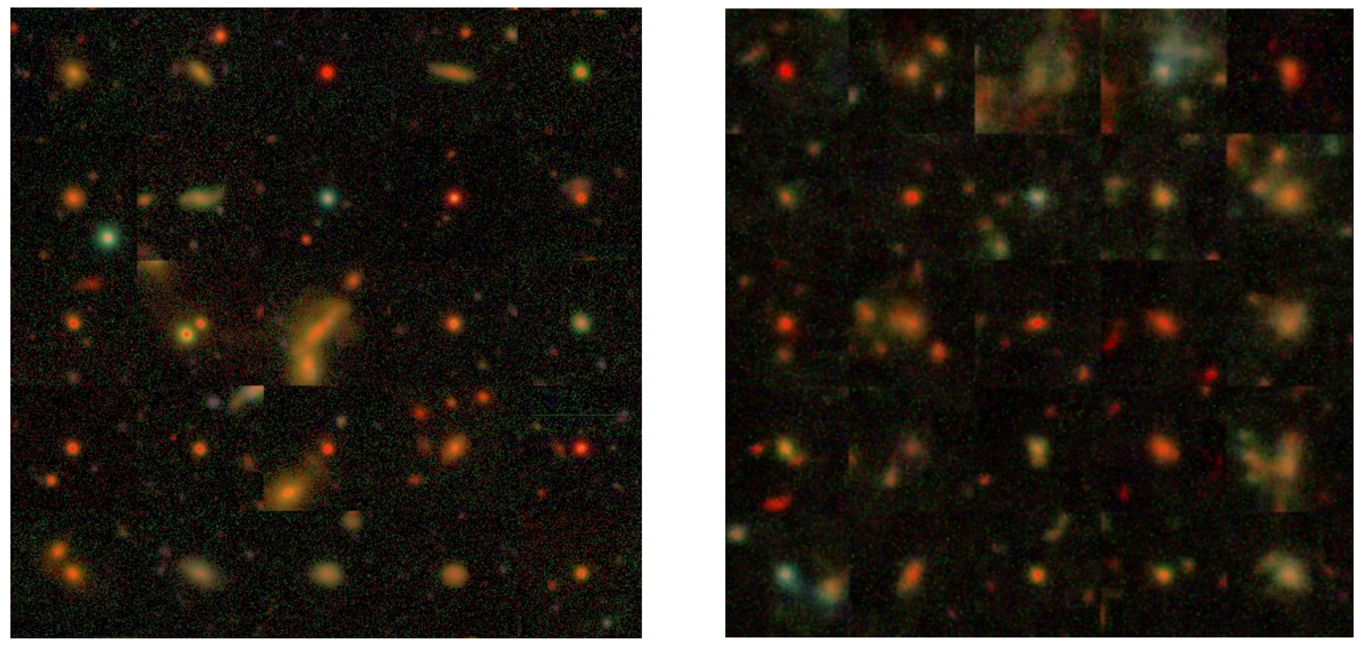
Marc Huertas-Company, Nesar Ramachandra, Francois Lanusse, Alexie Leauthaud, Yifei Luo, Song Huang

NEW YORK UNIVERSITY

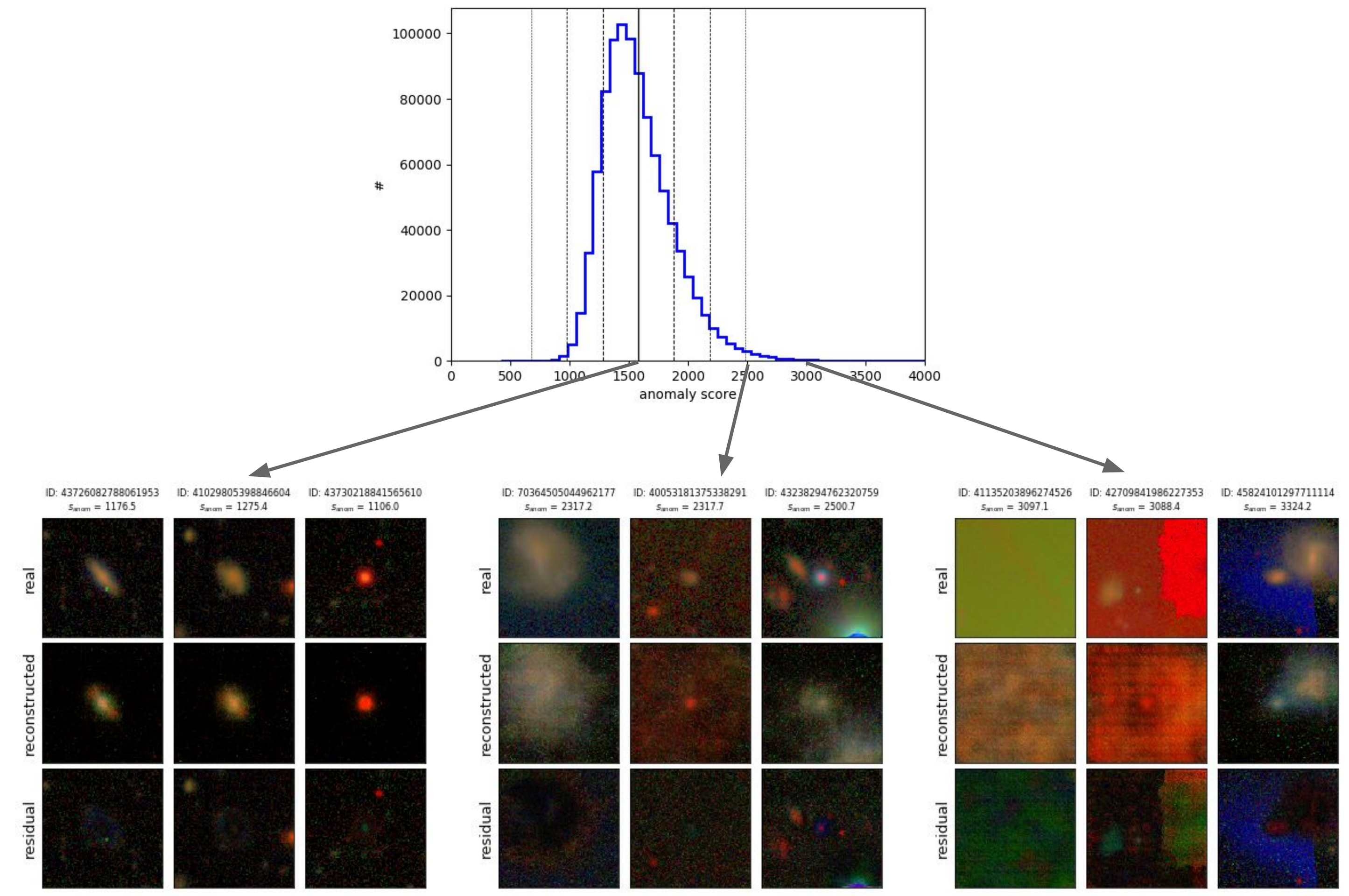
KAVLI SUMMER PROGRAM IN ASTROPHYSICS

We train a Wasserstein generative adversarial network (WGAN) on a set of ~1 million galaxies from the Hyper Suprime-Cam galaxy survey on the Subaru Telescope in Hawai'i.

The WGAN learns the data distribution and can generate new, similar-looking galaxies. *Can you tell which set is real and which is generated by the WGAN?*

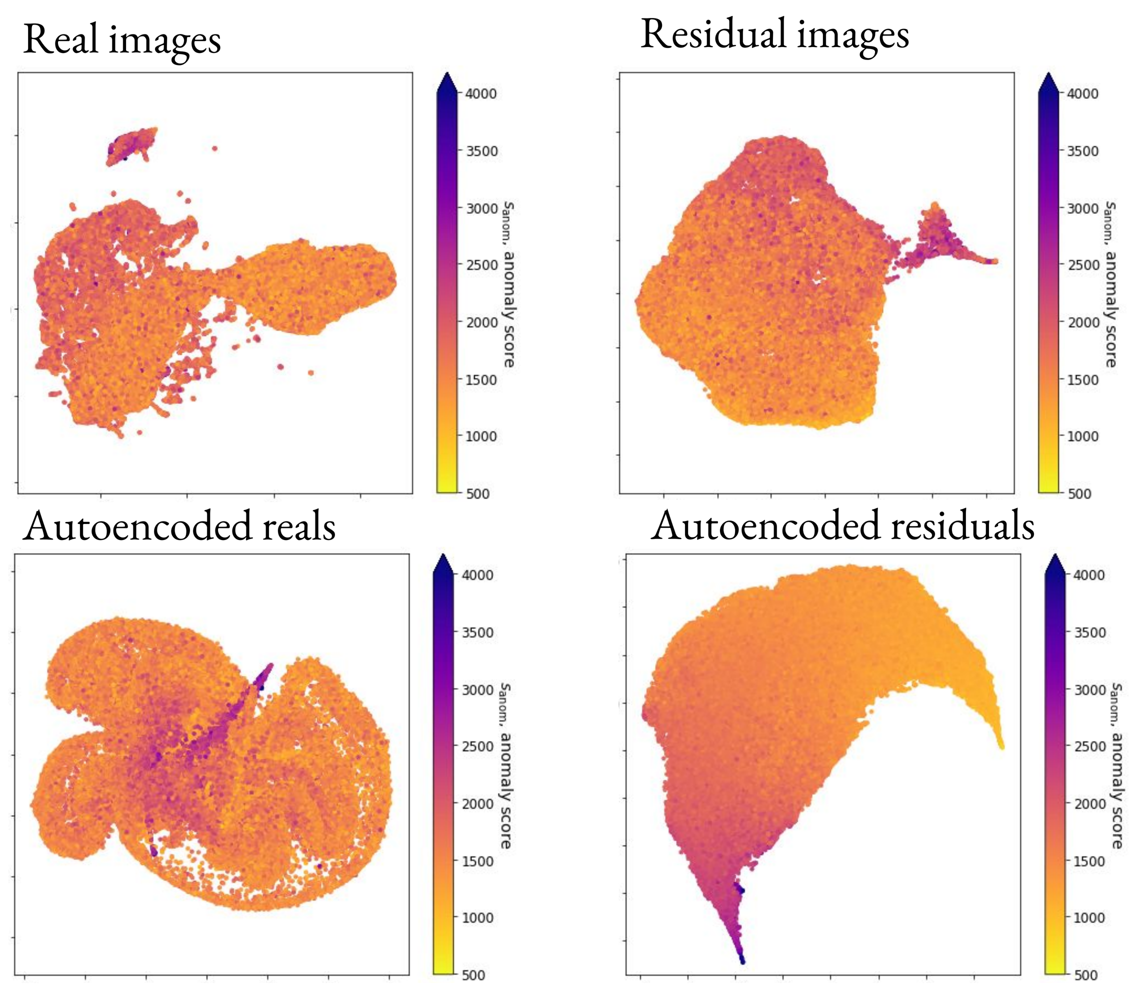


The images that the WGAN can reconstruct well are more typical, while images it cannot model are more *anomalous* with respect to the data. We assign "anomaly scores" based on this.



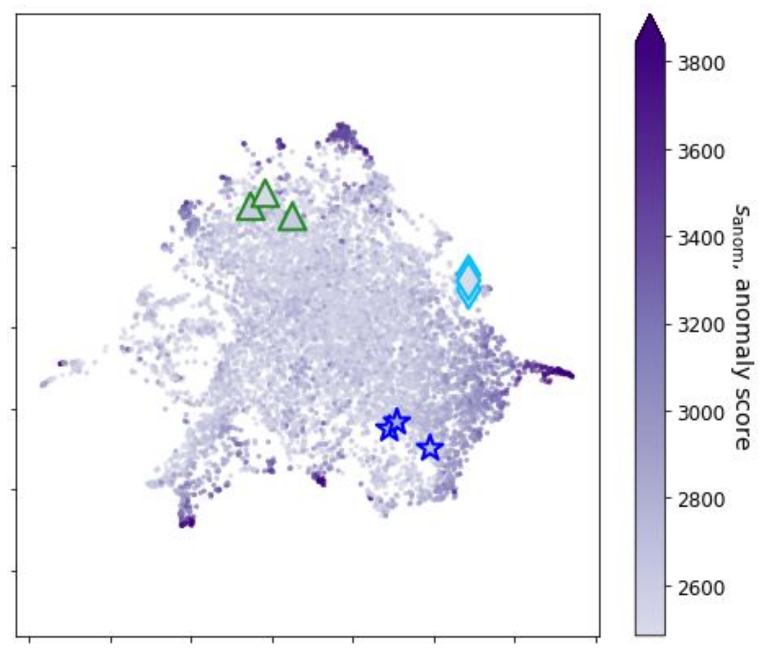
images near the mean ~3σ above the mean ~5σ above the mean

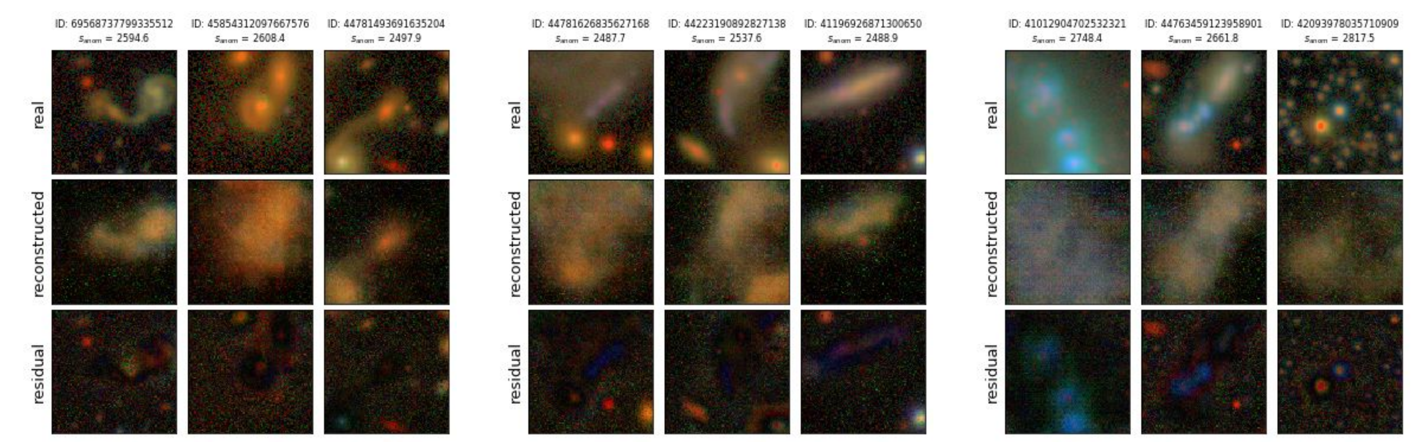
We train a convolutional autoencoder (CAE) to compress the images. The autoencoded residual between the image and the WGAN's best reconstruction correlates strongly with anomaly score, as they contain information about the anomalous features. See UMAP embeddings on...



We characterize the images with anomaly score >3σ using a UMAP on their autoencoded residuals. Similar anomalies cluster together in UMAP space..

We find scientifically interesting anomalies of many types:





△ merging galaxies ◇ purple active regions ☆ blue star formation

The astronomical surveys of the coming decade will image *billions* of objects.

We need unsupervised, scalable methods for finding interesting anomalous objects.

We show that **generative adversarial networks**, combined with convolutional autoencoders, can find scientifically interesting galaxies—like this one.

Find more yourself at https://weirdgalaxi.es!