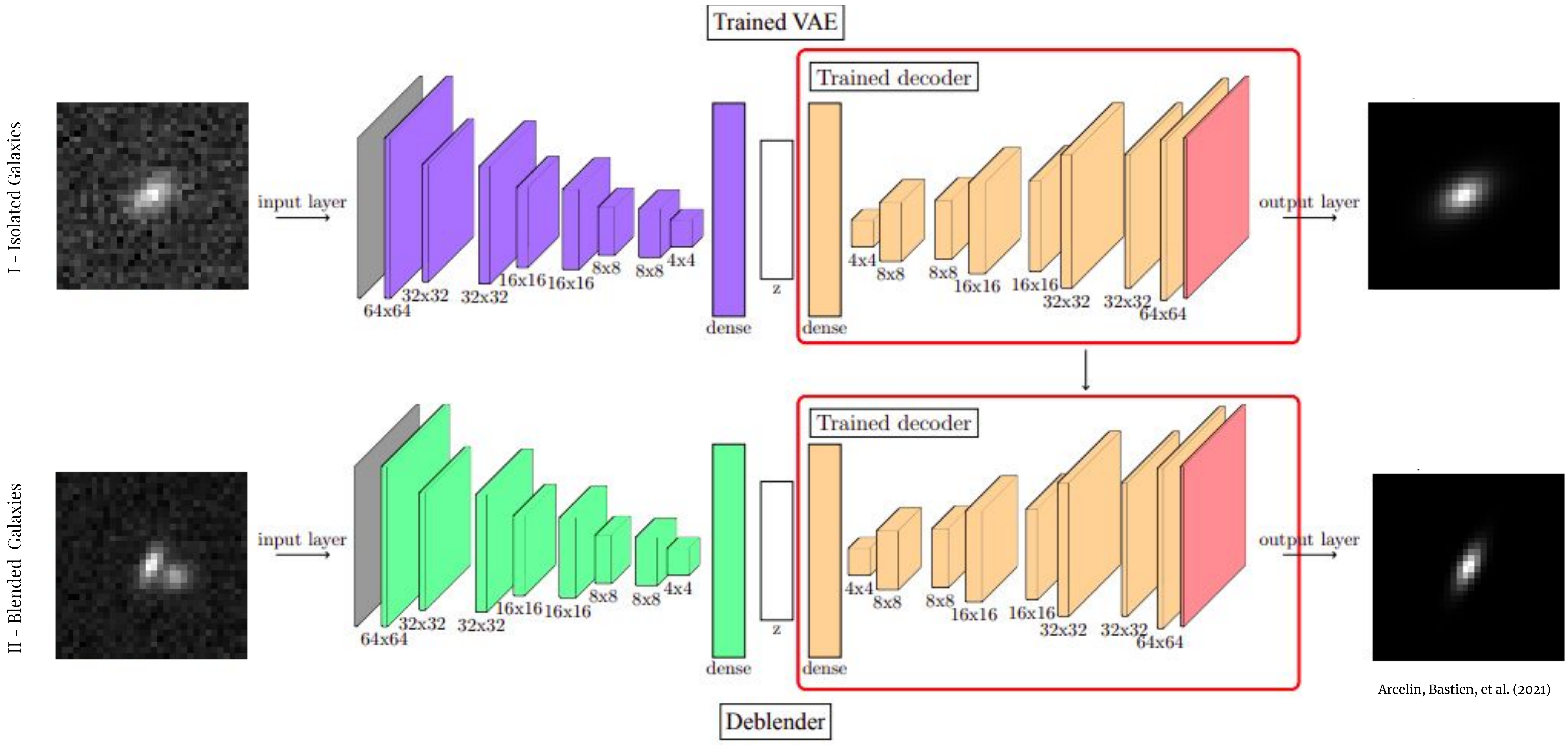
Reconstructing blended galaxies with Machine Learning

Lavanya Nemani,[1,2] Adriano Fontana[1] and Emiliano Merlin[1]
[1]Istituto Nazionale di Astrofisica - Osservatorio Astronomico di Roma, [2]Sapienza Università di Roma

MACHINE LEARNING FOR ASTROPHYSICS
CATANIA, 30 MAY - 1 JUNE, 2022

Galaxy blending is a confusion effect created by the projection of photons from galaxies on the same line of sight to the sky 2D plane. For standard **aperture photometry deblending is crucial**. The current standard deblending algorithms like **SExtractor** are based on threshold methods that assign each pixel to a single object. Instead we use two distinct Variational Autoencoders (VAEs) to deblend galaxies: One which learns how to reconstruct galaxy light profiles in isolation, and another one which uses the trained part of the first network to actually deblend overlapping pairs reconstructing their individual light profiles.The main focus of our work is to obtain **accurate flux and morphology estimates for blended objects** and clean **light profiles to be used as priors for template fitting codes**. With our current best approach we are able to retrieve the original flux within **10% for 1 sigma (whereas SExtractor is within 16% for 1 sigma).**



Arcelin, Bastien, et al. (2021)

Data Generation

Data generation with EGG [C. Schreiber et al. (2017)] which is a suite of tools that can generate mock galaxy catalogs with realistic positions, morphologies and fluxes from the far-ultraviolet to the far-infrared. Galaxy images are produced with GalSim [Rowe, Barnaby TP, et al. (2015)] which is an image simulation toolkit.

For the first VAE, we use the stamps as is and add Gaussian Noise (given by the limiting magnitude of EUCLID VIS band) to the input images.

For the second VAE, we artificially blend objects by superimposing a galaxy on another within a certain annulus. The pixel values are added in the regions of overlap.

Training

Each dataset is split into two for training and validation purposes with a 80/20 ratio. The images are normalized within bounds of [0,1].

The loss of the model is sum of two loss: reconstruction loss and kl-divergence loss. The training stops when the validation loss starts to diverge from the training loss with a early stopping condition with patience of 7.

Metrics like cosine distance and MSE are also used to evaluate the performance of the model. Finally, a proxy of the flux of the galaxy, given by sum of pixels of the image is used to estimate the errors of prediction of the blended objects. Examples of de-blendning can be seen below: