# Reconstructing blended galaxies with Machine Learning

Lavanya Nemani*[1], Emiliano Merlin[1], and Adriano Fontana[1]

[1]INAF - Osservatorio Astronomico di Roma – Italy

**Abstract**

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 (Dawson & Schneider 2014). The upcoming deep extragalactic surveys like LSST and EUCLID expect to see a blending fraction of up to 50% in the densest regions (Reiman & Göhre 2018). For standard aperture photometry and for more complex techniques such as psf-fitting and template-fitting algorithms, de-blending, the process of reconstructing the individual light profiles from blended sources, becomes crucial. The current standard de-blending algorithms (e.g. SExtractor, Bertin & Arnouts 1996) are based on threshold methods that simply assign each pixel to a single object, often failing to correctly take into account the real properties of the blended galaxies. With the advent of Machine Learning and Computer Vision in Astronomy we want to explore an unbiased and more accurate method of reconstructing individual light profiles using generative models.
Variational Auto-Encoders (Kingma & Welling 2013), VAE, are a type of probabilistic generative models that consist of two parts: In the encoder part the model learns to reduce the high-dimensional input to an encoded representation, and in the decoder part it learns how to reconstruct the input from the lower dimensional representation (called the bottleneck). Two distinct networks are needed to deblend galaxies using VAEs: 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 (e.g. Arcelin et al 2020).

In our work we simulate stamps of galaxy images as expected from the EUCLID survey in the VIS band, to test the results of using this ML technique for deblending as opposed to standard deblending methods like SExtractor. The galaxy images are built starting from a mock input catalogue created using EGG, a code that can generate mock galaxy catalogs with realistic positions, morphologies and fluxes (Schreiber et al 2017); the catalogue either feeds the image simulation toolkit GalSim (Rowe et al 2014) to generate analytical double-Sersic profiles, or it is used in parallel with autoencoders on HST images to obtain "euclidized" realistic profiles (Euclid Morphology Challenge, Bretonniere et al. in preparation).

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 (e.g. T-PHOT, Merlin et al 2015). We use several metrics to test our performance of the VAE algorithm comparing the reconstructed light profile of each simulated galaxy to the true input one, such as mean-square error and cosine distance. We also compare fluxes obtained on predicted images

*Speaker

with the input catalogue fluxes. 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). In the future, we want to optimize the network architecture used for training an ML algorithm to perform deblending by hyper-parameter training and including state-of-the-art ML architecture blocks like VGG (Simonyan et al 2014), Inception-Residual (Alom et al 2017), ResNet (Kaiming et al 2015).
Video: https://youtu.be/SnTG_rZ3Qig

**Keywords:** Computer Vision, EUCLID, Deblending