RNAAS RESEARCH NOTES OF THE AAS

An Evaluation of Autoencoder Based Models for Galaxy Deblending

To cite this article: Saahir Narang *et al* 2025 *Res. Notes AAS* **9** 237

Draft version August 28, 2025
Typeset using LaTeX default style in AASTeX631

An Evaluation of Autoencoder Based Models for Galaxy Deblending

Saahir Narang,[1] Sanika Nandpure,[1] and Trang Hoang[1]

[1] The University of Texas at Austin

ABSTRACT

Galaxy deblending is a critical challenge in astronomical surveys, particularly in large-scale surveys where overlapping galaxies complicate accurate observation and analysis. This paper evaluates autoencoder-based deep learning models for galaxy deblending, comparing convolutional autoencoders (CAE), Variational Autoencoders (VAE), and Variational Autoencoder-Generative Adversarial Networks (VAE-GAN). We find that CAEs have the best performance. We discuss the performance of the models as well as possible areas of improvements.

1. INTRODUCTION

When conducting large surveys of the sky and observing thousands of celestial bodies, galaxies within the same line of sight often overlap with each other. The sheer amount of galaxy data collected means that these overlaps occur fairly frequently, muddling data collection. We will refer to this phenomenon as "galaxy blending" (Melchior et. al 2021). The ability to disentangle these overlapping sources is essential for precise photometric and astrometric measurements.

In this paper, we will evaluate several autoencoder-based deep learning models for galaxy deblending, including convolutional autoencoders (CAE), Variational Autoencoders (VAE), and Variational Autoencoder- Generative Adversarial Networks (VAE-GAN).

Let us first review the models used. All three of our models were made using the PyTorch deep learning library (Paszke et al., 2019). Convolutional autoencoders compress images into a latent space via an encoder and reconstruct them with a decoder (Zhang 2018). We use a 4-layer convolutional encoder-decoder architecture with symmetric downsampling and upsampling. The encoder maps 64×64 RGB images through 3×3 stride-2 convolutions with channel sizes $3 \to 64 \to 128 \to 256 \to 512$ and standard ReLU activations.

Variational autoencoders add a probabilistic bottleneck, enabling greater variability in reconstructions (Kingma & Welling 2013). We implement a variational autoencoder with a 4-layer convolution-based encoder ($3 \to 64 \to 128 \to 256$, 5×5 stride-2) projecting to a fully connected layer of size 2048, from which $\mu, \log \sigma^2 \in \mathbb{R}^d$ are derived. The decoder maps latent vectors $z \in \mathbb{R}^d$ to $8 \times 8 \times 256$ and reconstructs images using three transposed convolutions ($256 \to 128 \to 32 \to 3$).

VAE-GANs build upon VAEs' efficient latent representations with Generative Adversarial Network (GAN) based discriminators that incentivize higher-quality samples (Larsen et al. 2016). Our VAE-GAN implementation augments our existing VAE architecture; the added discriminator applies four convolutions ($3 \to 32 \to 128 \to 256 \to 256$) followed by a fully connected projection ($256 \times 8 \times 8 \to 512 \to 1$).

2. DATASET

In order to train our models, we needed a large dataset of high-quality images so that the models could learn patterns that underly galaxy deblending tasks. Previous literature has tackled this problem in two ways: (i) artificially "blending" real galaxy images and (ii) generating synthetic images to train with (Reiman & Gohre 2019). Both of these methods have distinct advantages and disadvantages, but we opted to use the former. We decided to use a galaxy image dataset produced by the team behind the Galaxy Zoo project for a Kaggle challenge on galaxy classification (Dieleman et al. 2015)

To artificially blend images, we followed a process similar to that used in the Reiman and Gohre's paper on galaxy deblending (Reiman & Gohre 2019). To create a set of galaxy images, we first load two random batches of 128 images from the image dataset. We then "perturb" all of the images in the second batch by rotating and translating the images. The images are rotated by 1 radian and are translated along the x- and y-axes randomly by 5 to 10 pixels. This is done so that the final blend more accurately reflects real-world circumstances as galaxies don't tend to stack on top of each other exactly centered. Finally, we stack the images from both batches on top of each other by taking the highest RGB value for each pixel among each set of two images. This method of taking the highest RGB value was used by Reiman and Gohre to prevent diluting faint features since the majority of pixel values in the images are zero or close to zero. Given this data, the goal of our models is to reconstruct the images in the first batch - i.e. the galaxy the image is centered on. The final training dataset has 70k images.

3. RESULTS

We evaluated all three of the models we built using three metrics: Mean Squared Error (MSE), Peak signal-to-noise ratio (PSNR), and Structural similarity index measure (SSIM). These three values are standard metrics in image processing to compare the similarity of two images. A lower MSE indicates greater similarity while a higher PSNR or SSMI indicate a higher quality reconstruction.

Our results are summarized as follows: the CAE has a MSE, PSNR (in dB) and SSIM of 8.117E-4, 3.092E1 dB, and 8.067E-1, respectively; the VAE has values of 2.446E-3, 2.615E1 dB, and 5.955E-1, respectively; the VAE-GAN has values of 1.963E-3, 2.708E1 dB, and 6.308E-1, respectively. Of the three models, the CAE had the best performance, followed by the VAE-GAN, and then the VAE. Despite the strong performance of the CAE, the results also highlight several issues, particularly with maintaining the structural integrity of spiral galaxies in the reconstructed image. Figure 1 depicts these overall results. While most galaxies are reconstructed accurately, in a few extreme cases, we observed that the CAE would blur the arms of a spiral so that it appeared as an elliptical galaxy (see Sample 4 in Figure 1). As for the images generated by the VAE and VAE-GAN, they were noisier and less well-defined.

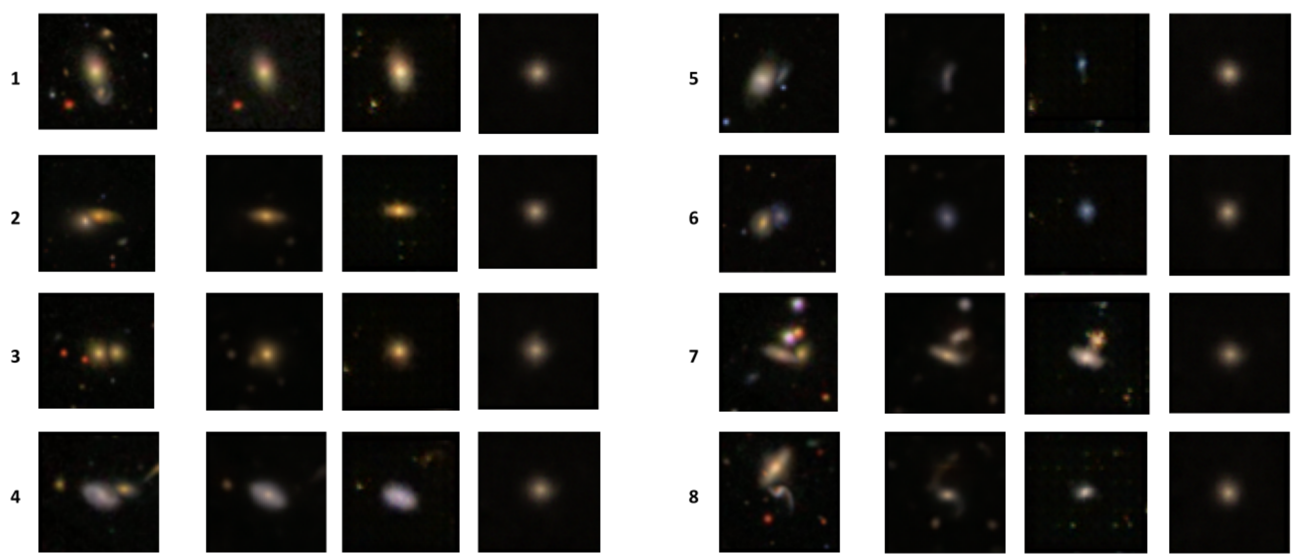

Figure 1. 8 samples of blended and deblended images. For each sample, left: blended input, right group: CAE, VAE-GAN, VAE outputs in that order.

4. CONCLUSION

This paper presented and evaluated autoencoder-based approaches for galaxy deblending. While past literature has used deep learning techniques including GANs, CNNs, and VAEs in isolation, this approach includes CAEs and hybrid models (VAE-GANs) and thus explores techniques that have not yet been used for galaxy deblending.

There is much work to be done in continuing to develop autoencoders as an effective means for galaxy deblending. In the future, we plan on building custom loss functions that account for astronomical parameters such as flux, ellipticity, and the point spread function (PSF) to maintain the structural integrity of the galaxy.

As mentioned, the VAE-GAN produced very blurry output images. This is likely because the decoder samples from a continuous latent space to reconstruct the output, meaning that the sampled features are not guaranteed to be accurate. One possible solution is to restructure the reconstruction term of VAE loss function such that it explicitly penalizes the generation of blurry images. This approach was proposed by Bredell et al. and showed promising results (Bredell et al. 2023). Another approach is to change the structure of the latent space itself; a variation of VAEs called Vector-Quantized Variational Autoencoders (VQ-VAEs) implements this. Instead of following a probabilistic distribution, the latent space of VQ-VAEs consists of discrete elements which can be optimized to produce sharper results. A possible solution could thus involve implementing a hybrid model that combines the discrete latent space of the VQ-VAE with the high-quality image generation ability of the GAN.

5. ACKNOWLEDGEMENTS

We would like to thank Dr. Shyamal Mitra from the Geometry of Space research group at The University of Texas at Austin for his support and guidance throughout our work.

REFERENCES

Arcelin, B., Doux, C., Aubourg, E., & Roucelle, C. 2020, Monthly Notices of the Royal Astronomical Society, 500, 531–547, doi: 10.1093/mnras/staa3062

Bredell, G., Flouris, K., Chaitanya, K., Erdil, E., & Konukoglu, E. 2023, Explicitly Minimizing the Blur Error of Variational Autoencoders. https://arxiv.org/abs/2304.05939

Burke, C. J., Aleo, P. D., Chen, Y.-C., et al. 2019, Monthly Notices of the Royal Astronomical Society, 490, 3952, doi: 10.1093/mnras/stz2845

Dieleman, S., Willett, K. W., & Dambre, J. 2015, Monthly Notices of the Royal Astronomical Society, 450, 1441, doi: 10.1093/mnras/stv632

Jarvis, J. F., & Tyson, J. A. 1981, The Astronomical Journal, 86, 476, doi: 10.1086/112907

Kingma, D. P., & Welling, M. 2013, Auto-Encoding Variational Bayes. https://arxiv.org/abs/1312.6114

Larsen, A. B. L., Sønderby, S. K., Larochelle, H., & Winther, O. 2016, in Proceedings of Machine Learning Research, Vol. 48, Proceedings of The 33rd International Conference on Machine Learning, ed. M. F. Balcan & K. Q. Weinberger (New York, New York, USA: PMLR), 1558–1566. https://proceedings.mlr.press/v48/larsen16.html

Melchior, P., Joseph, R., Sanchez, J., MacCrann, N., & Gruen, D. 2021, Nature Reviews Physics, 3, 712, doi: 10.1038/s42254-021-00353-y

Reiman, D. M., & Göhre, B. E. 2019, Monthly Notices of the Royal Astronomical Society, 485, 2617–2627, doi: 10.1093/mnras/stz575

Zhang, Y. 2018, ICONIP17-DCEC. https://api.semanticscholar.org/CorpusID:209442203