Spectral Super Resolution with Cross-Modal Diffusion Models

Aryana Haghjoo

February 5, 2025

1. Background and Motivation

The study of galaxies relies on both photometric images and spectroscopic data, which provide information such as redshift, morphology, and metallicity. With spectroscopy remaining an expensive and time-consuming task, spectroscopic surveys are limited in coverage compared to photometric surveys. Consequently, spectral data for many galaxies are incomplete or unavailable. Furthermore, low-resolution spectra from space-based surveys lack the detailed spectral features necessary for precise measurements. As an example, the line blending issue (e.g., NII/Hα) introduces systematic bias in redshift estimation and consequently inference of physical parameters [1].

Inspired by Parker et al. (2024), where the authors use DESI data to train a transformer-based encoder [2], we propose a cross-modal foundation model that leverages diffusion models as an encoder. This model creates a shared embedding space between images and spectra. It enables multiple **downstream tasks**, including:

- **Generating spectra from images** for galaxies lacking spectroscopic observations
- **Enhancing the resolution of spectral data** to de-blend atomic lines
- **Enabling other cross-modal applications** uch as anomaly detection.

By leveraging the generative power of diffusion models, this approach aims to create a unified representation that captures complex relationships between imaging and spectral data.

In in a similar study by Doorenbos et al. (2022), authors introduced a multimodal conditional diffusion model to generate galaxy spectra from photometric images, validating the approach using SDSS data and showcasing its feasibility [3]. Building on this foundation, Doorenbos et al. (2024) extended the methodology to not only generate spectra but also infer galaxy properties, such as stellar velocity dispersion, metallicity, and the presence of an AGN [4]. The code from these studies is public and available at this link.

2. Data Sets and Training Strategy

With Euclid data yet to be released, we will use the following datasets for training the network:

- Low-resolution: **3DHST GRISM spectra**
- High-resolution:**Keck MOSDEF 1D spectra, KBSS (MOSFIRE)**

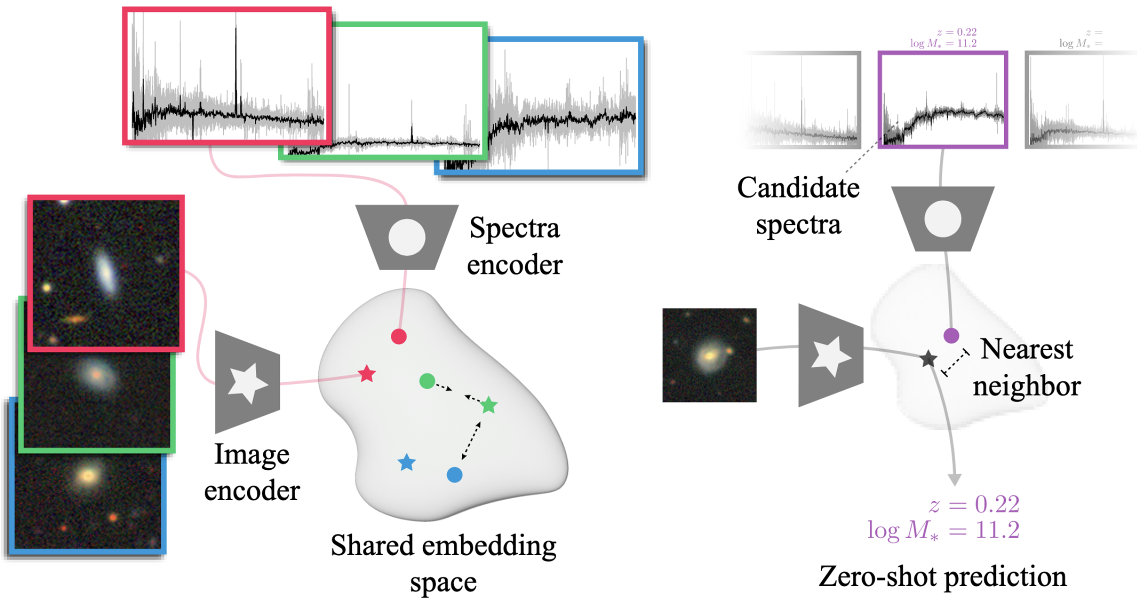

Figure 1: AstroCLIP's cross-modal training strategy involves two steps. First, galaxy images and spectra are separately encoded using a transformer-based self-supervised network. Then, a cross-modal contrastive loss aligns these encoders, enabling meaningful comparisons and preserving physical information for various downstream tasks. We aim to do a similar job but our approach is based on diffusion models using a different data set as our training set. Figure from [2].

Following a similar approach to the method described in Figure 1, we will employ two separate diffusion-based encoders for images and spectra. These encoders will then be aligned using a contrastive loss function. This process will create a lower-dimension shared embedding space. A fraction of the data not included in the training set will be used to evaluate the network, ensuring the accuracy of the encoding and decoding processes.

It is expected that the 3DHST GRISM spectra will contain a significant level of noise. If the desired results are not achieved with this dataset, we can resort to simulated data. The CLOUDY simulation might be a suitable alternative.

3. Future Applications

Once the **Euclid Q1 data** is released, we will use it as our low-resolution dataset to generate high-resolution spectra comparable to those expected from the **Roman** telescope. This will help mitigate the issue of NII/Hα blending.

Our results can be compared with those of Faisst et al., who modeled the [N II]/Hα flux ratio as a function of stellar mass and redshift [5].

References

[1] Daniel Martens et al. "Effects of [N ii] and H α line blending on the WFIRST Galaxy redshift survey". In: *Monthly Notices of the Royal Astronomical Society* 485.1 (2019), pp. 211–228.

[2] Francois Lanusse et al. "Astroclip: Cross-modal pre-training for astronomical foundation models". In: *NeurIPS 2023 AI for Science Workshop*. 2023.

[3] Lars Doorenbos et al. "Generating astronomical spectra from photometry with conditional diffusion models". In: *arXiv preprint arXiv:2211.05556* (2022).

[4] Lars Doorenbos et al. "Galaxy spectroscopy without spectra: Galaxy properties from photometric images with conditional diffusion models". In: *The Astrophysical Journal* 977.1 (2024), p. 131.

[5] Andreas L. Faisst et al. "Empirical Modeling of the Redshift Evolution of the [N II]/Hα Ratio for Galaxy Redshift Surveys". In: *The Astrophysical Journal* 855.2 (2018), p. 132.