# Learning to See Sharper: A Physics-Informed Artificial Intelligence Framework for Super-Resolving Galaxy Spectra

**Aryana Haghjoo**[1] **Shoubaneh Hemmati**[2]

**Bahram Mobasher**[1]

[1]Dept. of Physics & Astronomy, UC Riverside [2]IPAC, California Institute of Technology

ahagh010@ucr.edu

## Abstract

The scientific information recoverable from galaxy spectra is fundamentally set by the resolution, yet assembling large high-resolution samples remains observationally expensive. We present a three-stage physics-informed deep-learning framework for spectral super-resolution that enhances *James Webb Space Telescope* (JWST)/Near Infrared Spectrograph (NIRSpec) low-resolution prism spectra ($\mathcal{R}\sim100$) by a factor of $\sim$10 in resolving power ($\mathcal{R}\sim1000$), recovering medium-resolution detail from prism-only observations. Trained on 1,187 paired *JWST Advanced Deep Extragalactic Survey* (JADES) observations, our model achieves noise-limited residuals across most of the 1–5 $\mu$m range and systematically improves the signal-to-noise (S/N) of key diagnostic emission lines, such as [O II], H$\beta$, [O III], and H$\alpha$, often by factors of several. The architecture embeds spectroscopic physics directly: a dedicated redshift-inference head provides global spectral context, while a line-token self-attention branch parameterizes individual emission features as Gaussian profiles constrained by physical inter-line relationships. The framework can be extended to forthcoming wide-field grism surveys (*Euclid*, *Roman Space Telescope*), enabling population-level diagnostics across millions of galaxies otherwise inaccessible at low resolution.

## 1 Introduction

The physical information encoded in galaxy spectra is gated by observed resolution. From redshifts and star-formation rates to gas-phase metallicities and active galactic nuclei (AGN) diagnostics, the richest spectroscopic diagnostics require well-resolved line features. Emission lines close in wavelength such as H$\alpha$/[N II], the [O III] $\lambda\lambda4959, 5007$ doublet, blend irretrievably at low resolution, biasing every downstream physical quantity extraction (Faisst et al., 2018; Agostino et al., 2023).

Modern wide-area missions face an inherent tension between survey volume and spectral quality. Space observatories such as *Euclid* and the *Roman Space Telescope* will deliver grism spectroscopy for tens of millions of galaxies at $\mathcal{R}\sim100$–300, far below what resolves fine diagnostic lines at many redshifts (Laureijs et al., 2010; Spergel et al., 2013). Targeted programs such as JADES with JWST/NIRSpec provide exquisite medium-resolution data but cover far fewer targets.

One way to circumvent this trade-off is to leverage the small but growing number of galaxies observed at both low and high spectral resolution, and use machine learning to learn the mapping between the two, enabling the resolution enhancement of the much larger low-resolution samples. Machine learning has found broad application in astronomical data analysis (Hemmati et al., 2022; Sanjaripour et al., 2024; Parker et al., 2024), yet spectral resolution enhancement remains largely unexplored. The model is trained on paired spectra from the *JWST Advanced Deep Extragalactic Survey* (JADES;

Conference on Physics and AI at Stanford University (PAI 2026).

Eisenstein et al. 2023), where low-resolution observations are matched with medium-resolution spectra of the same galaxies.

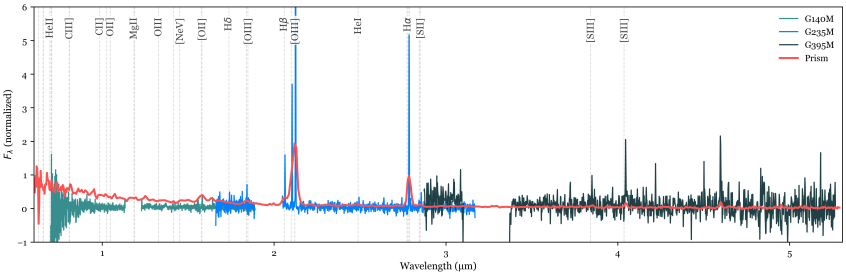


Figure 1: Example paired JWST/NIRSpec spectra for a representative JADES galaxy ($z = 3.235$). The prism (red) covers 0.6–5.3 $\mu$m at $\mathcal{R} \sim 100$; the three medium-resolution ($\mathcal{R} \sim 1000$) grating segments (G140M, G235M, G395M; green, blue, black) are overplotted. Vertical dashed lines mark key emission lines at the galaxy's redshift. The prism blends features that the gratings can resolve.

## 2 Data: Paired JWST/NIRSpec JADES Spectra

Our training dataset is drawn from the JADES dataset Eisenstein et al. 2023; Bunker et al. 2024; D'Eugenio et al. 2025; Curtis-Lake et al. 2026), conducted with the NIRSpec instrument. JADES observes the same galaxies in both the prism disperser ($\mathcal{R} \sim 100$, 0.6–5.3 $\mu$m) and three medium-resolution grating modes (G140M, G235M, G395M; combined 1–5.5 $\mu$m, $\mathcal{R} \sim 1000$), providing naturally co-observed paired spectra. An example of the prism and medium-resolution spectra for a representative JADES galaxy is shown in Figure 1.

**Pre-processing.** Sources observed in both modes are cross-matched by sky position (separation $< 0.2''$) and field name. The three grating spectra are merged into a single continuous reference by minimizing root mean square (RMS) flux differences in overlap regions and applying consistent flux scale factors. Both prism and stitched grating spectra are placed on a common wavelength grid defined by the native grating sampling; the prism is up-sampled via cubic interpolation, which preserves broadband shape without introducing artificial structure. Galaxies are retained only if their medium-resolution spectrum shows $\mathrm{S/N} > 5$ in at least one of H$\alpha$, [O II], or [O III], yielding **1,187 paired spectra**.

**Augmentation.** Each pair generates 20 stochastic realizations via Gaussian redshift offsets ($\sigma_z = 0.3$) and 10% flux perturbations, yielding 24,927 augmented pairs that promote shift-robust, noise-tolerant learning.

## 3 Physics-Informed Super-Resolution Framework

Our pipeline comprises three sequentially trained components; Figure 2 shows the overall design.

**Stage 1 — Coarse Super-Resolution (SR1).** SR1 is a 1D residual convolutional neural network that maps the prism spectrum $x_{\mathrm{lr}}$ to an initial super-resolved estimate $x_{\mathrm{sr1}}$ and a predicted log-variance $\log \sigma^2_{\mathrm{sr1}}$. It is trained with a heteroscedastic negative log-likelihood loss that reflects the wavelength-dependent noise of the NIRSpec detector. The architecture uses $\sim 8$ residual blocks with small convolutional kernels and hidden dimension $\sim 128$, providing a physically plausible spectral backbone on which subsequent stages operate.

**Stage 2 — Redshift Inference (ZHead).** ZHead is a 1D CNN applied to the SR1 output that infers a redshift estimate $\hat{z}$ and its uncertainty $\log \sigma^2_{\hat{z}}$. Treating redshift as an explicit global latent parameter captures long-range spectral coherence that local convolutions cannot encode: line pairs relevant to one another (e.g., [O III] $\lambda\lambda 4959, 5007$ and H$\beta$) may be separated by hundreds of pixels in the observed frame. During SR2 training, a teacher-forcing schedule linearly decays the probability of substituting the catalog redshift for $\hat{z}$ from 0.8 to 0.2, transitioning the model from supervised to self-consistent redshift inference.

**Stage 3 — Physics-Informed Residual Refinement (SR2).** SR2 predicts a residual correction $\Delta x = \Delta x_{\mathrm{line}} + \Delta x_{\mathrm{CNN}}$ to the SR1 output via two parallel branches.

*Line-token self-attention branch.* For each of $K = 20$ known rest-frame emission lines, a local spectral window around the expected observed-frame wavelength at $\hat{z}$ is encoded by a small CNN into a compact feature vector; a learnable species embedding distinguishes morphologically similar lines. The $K$ tokens are processed by a Transformer encoder, enabling every line to exchange information

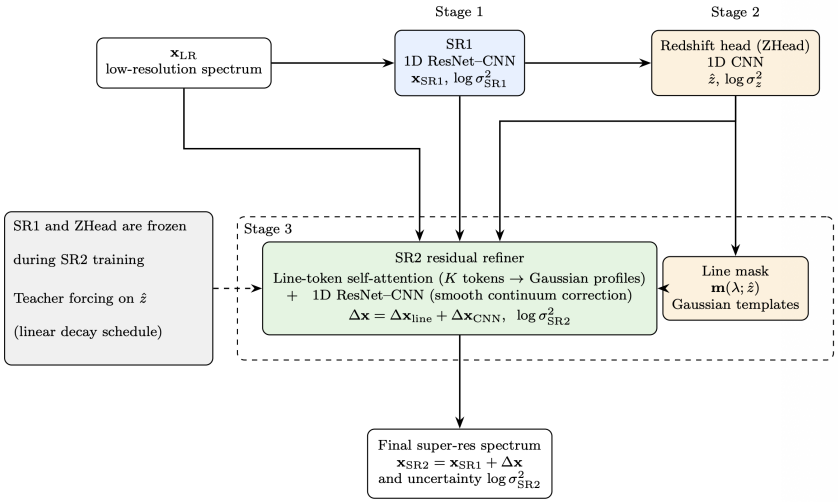


Figure 2: Three-stage super-resolution architecture. **SR1**: a 1D ResNet-CNN maps the prism spectrum to a coarse super-resolved estimate and predictive uncertainty. **ZHead**: a 1D CNN infers the galaxy redshift $\hat{z}$ from the SR1 output. **SR2**: a two-branch residual refiner conditioned on $\hat{z}$, combining a line-token self-attention branch (parametric Gaussian profiles per emission line) with a convolutional branch for smooth continuum correction.

globally. This explicitly encodes physical inter-line relations ($\approx 1{:}3$ [O III] doublet ratio, Balmer decrement) inaccessible to local convolutions. Each token decodes into a parametric Gaussian profile; a redshift-conditioned wavelength mask suppresses outputs where no feature is expected.

*Continuum correction branch.* A shallow ResNet-CNN with scaled skip connections ($\times 0.5$) predicts smooth continuum deviations $\Delta x_{\mathrm{CNN}}$, biased toward low-frequency adjustments. The two-branch design physically decomposes the galaxy spectrum into discrete emission features plus a slowly varying continuum.

**Hyperparameter Optimization.** We performed three sequential Bayesian sweeps using Weights & Biases (Biewald, 2020), with a total of $\sim 100{,}000$ trials, exploring learning rate, weight decay, dropout, hidden dimension, residual blocks, and activation per stage; SR2 additionally covered attention heads/layers, spectral window width, and L1 amplitude penalty. Optimal values across all stages consistently favor exponential linear unit (ELU) activations, learning rates $\sim 10^{-4}$, and $\sim 8$ residual blocks.

**Computational resources.** All stages were trained on a single **NVIDIA RTX 3090** (24 GB VRAM); final models require 2–4 hours per stage, with the full Bayesian sweep costing $\approx 400$ GPU-hours. The pipeline totals $\approx 2.5$M parameters; inference takes $< 1$ s per spectrum.

# 4 Results

We evaluate performance on the 20% held-out evaluation set.

**Spectral reconstruction.** Figure 3 shows example reconstructed spectra for two held-out galaxies. SR2 resolves the [O III] $\lambda\lambda 4959, 5007$ doublet into two peaks at the correct flux ratio ($\approx 1{:}3$) while the prism input shows only broad, blended features. Residual analysis confirms that SR2–high-resolution (HR) residuals are noise-limited: the robust scatter $\sigma_{\mathrm{resid}}(\lambda)$ converges to the median HR noise amplitude $\tilde{\sigma}_{\mathrm{HR}}(\lambda)$ above $\sim 2\,\mu$m.

**Emission-line S/N improvement.** For each galaxy and diagnostic line, we fit a Gaussian model $f(\lambda) = C_0 + C_1\lambda + A\exp[-(\lambda-\mu)^2/2\sigma^2]$ to the low-resolution (LR) and super-resolved (SR) spectra, computing $\mathrm{S/N} = A/\delta A$ from the fit covariance. Figure 4 shows that for all four key diagnostics, the majority of galaxies lie above the one-to-one relation, confirming systematic improvement. These gains translate into higher-fidelity emission line detection and thus more precise estimates of the physical characteristics of the galaxy.

**Redshift consistency as a downstream diagnostic.** Identical redshift-inference networks trained on LR, SR, and HR spectra show SR reduces the normalized median absolute deviation [NMAD; $|\Delta z|/(1+z)$] by $\approx 50\%$ relative to LR, cutting the outlier fraction from $\sim 16\%$ to $\sim 12\%$ (Figure 5).

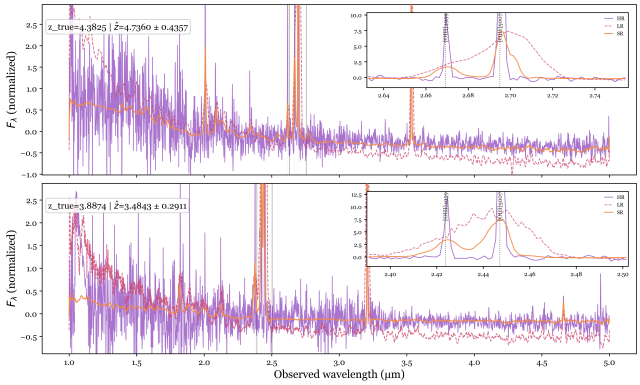


Figure 3: Example super-resolution outputs for two held-out JADES galaxies. Low-resolution prism input (red dashed), super-resolved SR2 output (orange), and medium-resolution reference (purple) over the full 1–5 $\mu$m range. Insets zoom into the [O III] $\lambda\lambda 4959, 5007$ showing deblending of features entirely unresolved at prism resolution.

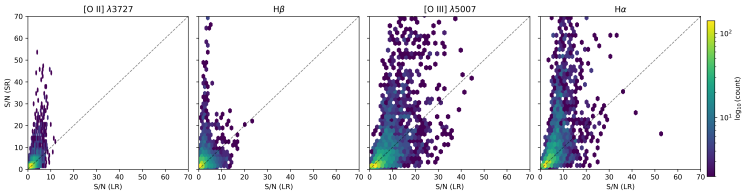


Figure 4: Emission-line S/N (LR vs. SR) for [O II], H$\beta$, [O III], and H$\alpha$. Points above the one-to-one line indicate improvement; color encodes $\log_{10}$ density.

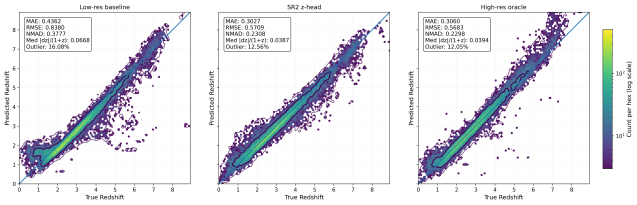


Figure 5: Predicted vs. true redshifts for LR (left), SR2 (middle), and HR (right) networks. SR2 substantially reduces scatter and outlier fraction, approaching the HR oracle. The true redshifts are obtained from the medium-resolution JADES spectral catalog.

## 5 Discussion and Conclusion

We have presented a physics-informed deep-learning framework that achieves noise-limited spectral super-resolution of JWST/NIRSpec galaxy spectra. By embedding spectroscopic domain knowledge (e.g., redshift inference, line-token self-attention, and a two-branch refinement stage) the model recovers spectroscopic information consistently with underlying physics. The downstream scientific impact is substantial: improved S/N enables higher-fidelity Baldwin-Phillips-Terlevich (BPT) diagramsBaldwin et al. 1981, star-formation rate densities, and gas metallicity gradients. The broad 0.6–5.3 $\mu$m coverage makes the framework applicable to *Euclid* and *Roman Space Telescope* grism surveys delivering spectra for tens of millions of galaxies. Future extensions include incorporating broadband photometry, expanding the training set through upcoming JWST programs, and applying the framework to absorption-line features. Physics-informed spectral super-resolution can bridge the resolution gap between current deep spectroscopic programs and forthcoming wide-area surveys, enabling population-level diagnostics at a fraction of the observational cost.

**Limitations.** The training set of 1,187 spectra introduces mild selection bias toward emission-line galaxies. Performance degrades at $\lesssim 1.5$ $\mu$m where both prism and HR data are observational noisy. Transfer to *Euclid* or *Roman* grism data will require domain adaptation for differences in resolution profile and detector noise.

## References


Agostino, C. J., Salim, S., Boquien, M., et al. 2023, Monthly Notices of the Royal Astronomical Society, 526, 4455

Baldwin, A., Phillips, M., & Terlevich, R. 1981, Publications of the Astronomical Society of the Pacific, 93, 817

Biewald, L. 2020, Experiment Tracking with Weights and Biases, `https://www.wandb.com/`

Bunker, A. J., Cameron, A. J., Curtis-Lake, E., et al. 2024, Astronomy & Astrophysics, 690, A288

Curtis-Lake, E., Cameron, A. J., Bunker, A. J., et al. 2026, Monthly Notices of the Royal Astronomical Society, stag836

D'Eugenio, F., Cameron, A. J., Scholtz, J., et al. 2025, The Astrophysical Journal Supplement Series, 277, 4

Eisenstein, D. J., Willott, C., Alberts, S., et al. 2023, arXiv preprint arXiv:2306.02465

Faisst, A. L., Masters, D., Wang, Y., et al. 2018, The Astrophysical Journal, 855, 132

Hemmati, S., Huff, E., Nayyeri, H., et al. 2022, The Astrophysical Journal, 941, 141

Laureijs, R. J., Duvet, L., Sanz, I. E., et al. 2010, in Space Telescopes and Instrumentation 2010: Optical, Infrared, and Millimeter Wave, Vol. 7731, SPIE, 453–458

Parker, L., Lanusse, F., Golkar, S., et al. 2024, Monthly Notices of the Royal Astronomical Society, 531, 4990

Sanjaripour, S., Hemmati, S., Mobasher, B., et al. 2024, The Astrophysical Journal, 977, 202

Spergel, D., Gehrels, N., Breckinridge, J., et al. 2013, arXiv preprint arXiv:1305.5425

# NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction accurately state our three main contributions: (1) the three-stage physics-informed architecture, (2) noise-limited spectral reconstruction on a 20% held-out evaluation set, and (3) systematic S/N improvements for four key diagnostic emission lines. All quantitative claims are supported by the results in Section 4.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: Section 5 discusses limitations including training-set selection bias toward emission-line galaxies, performance degradation at $\lesssim 1.5\,\mu$m, and expected domain shift when transferring to *Euclid* or *Roman* grism spectra.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [NA]

Justification: This paper does not include theoretical proofs. The contributions are empirical and architectural.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: Section 3 describes the full pipeline architecture, training procedure, loss functions, teacher-forcing schedule, and all optimal hyperparameter values. The JADES dataset is publicly available through the JADES team's data releases (Curtis-Lake et al., 2026).

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [No]

Justification: The JADES spectroscopic data are publicly available via the JADES data releases. Our model code will be made available upon acceptance. At submission time, reproducibility is supported by the architectural and training details in Section 3.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [Yes]

Justification: We report the 80/20 train/test split, the three-stage sequential training procedure, all hyperparameter values, the augmentation strategy, and the evaluation metrics (residual scatter, S/N ratios, NMAD) used throughout Section 4.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [Yes]

Justification: Reconstruction quality is quantified via the robust scatter $\sigma_{\mathrm{resid}}(\lambda)$ (MAD-based standard deviation) across the evaluation set, compared to the median HR noise $\tilde{\sigma}_{\mathrm{HR}}(\lambda)$. Redshift performance is quantified via NMAD and catastrophic outlier fraction. S/N scatter plots show the full distribution across $\sim 237$ evaluation galaxies.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)

- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

Answer: [Yes]

Justification: Section 3 reports that all experiments were run on a single NVIDIA GeForce RTX 3090 GPU (24 GB VRAM). Individual training runs complete in 1–2 hours; final models converge in 2–4 hours per stage. The full Bayesian hyperparameter optimization across all three stages required $\approx 400$ GPU-hours. Inference takes $< 1$ second per spectrum on the same hardware. This becomes signifant when the model is applied to survey-scale data in later stages.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
- The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
- The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

Answer: [Yes]

Justification: This work develops spectral analysis tools for astronomical observations. There are no foreseeable ethical concerns associated with this research domain.

Guidelines:

- The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [Yes]

Justification: Section 5 discusses the positive impact of enabling population-level galaxy diagnostics at reduced observational cost. This domain-specific astronomical tool has no identifiable negative societal impacts.

Guidelines:

- The answer NA means that there is no societal impact of the work performed.
- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification: This paper poses no such risks. The model operates on astronomical spectra and has no applicability to sensitive personal data or safety-critical systems.

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [Yes]

Justification: The JADES survey data are publicly released and cited throughout. PyTorch and Weights & Biases are open-source tools cited in Section 3. All other referenced software and datasets are properly attributed in the bibliography.

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.

- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [No]

    Justification: The trained model and inference code will be released upon acceptance. At submission time, the architecture and training procedure are fully documented in Section 3 to facilitate independent reimplementation.

    Guidelines:

    - The answer NA means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA]

    Justification: This paper does not involve crowdsourcing or research with human subjects. All data are astronomical observations from the JWST/NIRSpec instrument.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA]

    Justification: This paper does not involve crowdsourcing or research with human subjects. No IRB approval was required.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

Answer: [NA]

Justification: No LLMs were used as components of the core method. The architecture employs standard CNNs and Transformers trained from scratch on astronomical spectra.

Guidelines:

- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy (`https://neurips.cc/Conferences/2025/LLM`) for what should or should not be described.