Machine Learning Techniques for High-Redshift Galaxy Classification with JWST NIRCam Data

Rosa Roberts, Christopher J. Conselice

Jodrell Bank Centre for Astrophysics, University of Manchester, Manchester, M13 9PL, UK.

rosa.roberts@student.manchester.ac.uk, conselice@manchester.ac.uk

Introduction

The detection of very high-redshift galaxies provides vital insights into the early Universe, including the Epoch of Reionisation. The James Webb Space Telescope's (JWST) deep, multi-band near-infrared imaging enables such discoveries by capturing light from distant galaxies stretched by cosmic expansion. Traditional redshift (z) estimation faces growing limitations with large datasets and degeneracies between high-redshift galaxies and low-redshift contaminants like brown dwarfs. Machine learning offers a scalable alternative, capable of capturing complex patterns in imaging data, including spectral features like the **Lyman break**. This project investigates both supervised and semi-supervised approaches [1] to identify high-redshift candidates from JWST photometry and assess the performance of these methods for future surveys.

The Lyman Break

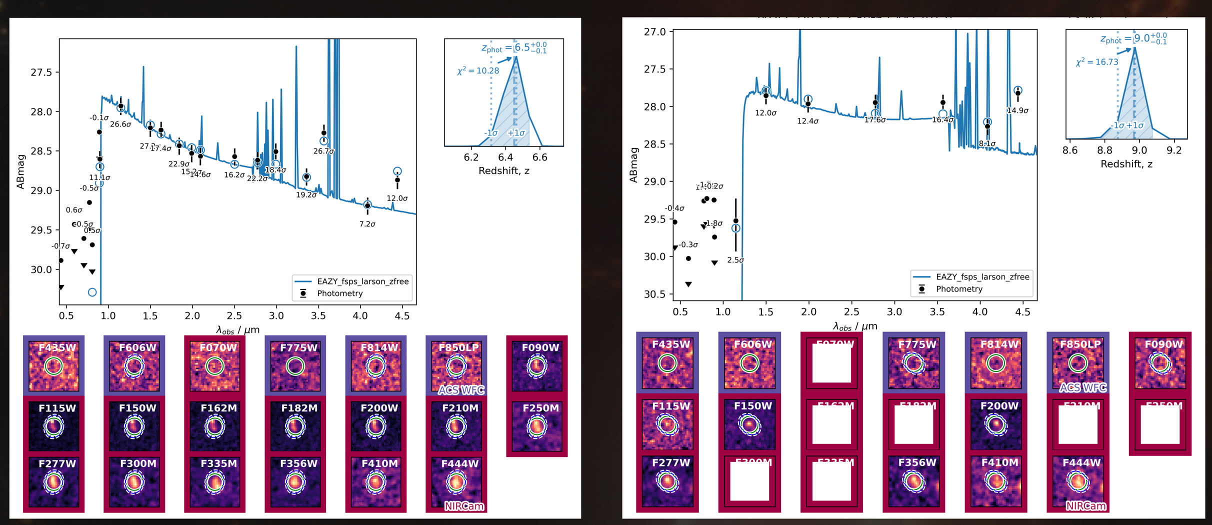


Figure 1: Galaxy SEDs at redshifts ~6.5 and ~9.0, illustrating the Lyman break shifting from ~0.9 µm to ~1.2 µm. NIRCam filter cutouts demonstrate how this shift across bands facilitates photometric redshift estimation. White boxes indicate missing filter coverage.

The **Lyman break** is a sharp drop in a galaxy's spectrum at 1216 Å (rest-frame) from hydrogen absorption. As redshift increases, it shifts to longer wavelengths:

$$\lambda_{obs} = 1216\,\text{Å} \times (1+z)\,.$$

This causes galaxies to "drop out" of bluer filters. NIRCam's 0.9–4.4 µm (F090W–F444W) coverage allows redshift estimation by tracking where the break appears across filters.

1. Supervised Pipeline: Data & Preprocessing

Using JADES DR3 GOODS-S [2] imaging in seven NIRCam filters, we select 85,420 sources with quality cuts on S/N, redshift PDFs, SED fits, and morphology [3][4]. From this, we build a balanced set of 722 low-z, 722 high-z galaxies, and 11 brown dwarfs (via atmospheric models) [5]. For each, we extract 64×64 cutouts, apply segmentation, normalisation, augmentation, and compute HOG features.

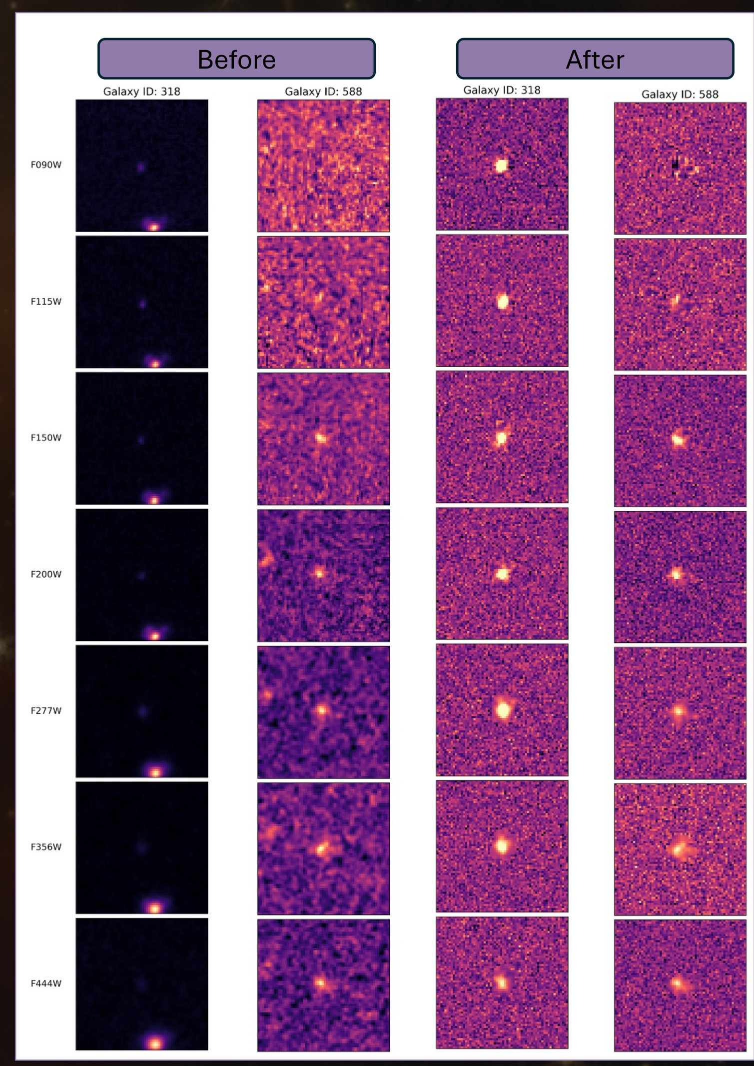


Figure 2. NIRCam cutouts of low-z (318) and high-z (588) galaxies before and after segmentation and flux normalisation. Galaxy 588 shows the typical dropout in F090W and F115W filters.

2. CNN Architecture

We use a ResNet-based CNN to classify 7-band JWST cutouts into low- z, high- z, or brown dwarf classes. The architecture separates spectral and spatial processing:

- **SpectralResidualBlock** captures inter-band features via 1×1 convolutions with dropout and batch norm.
- **SpatialResidualBlock** extracts spatial patterns using 3×3 convolutions, spatial dropout, and downsampling.
- **Channel Attention** (Squeeze-and-Excitation) reweights bands based on pooled statistics to boost key spectral signals.

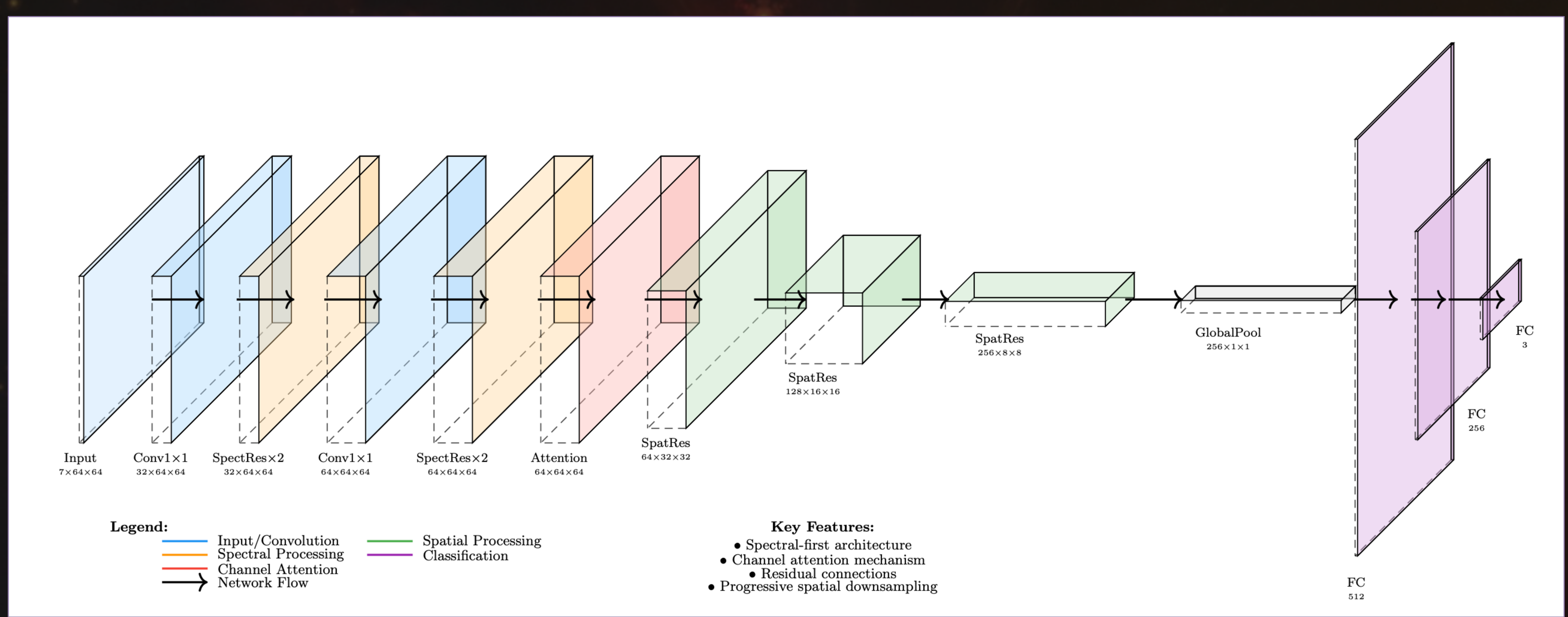
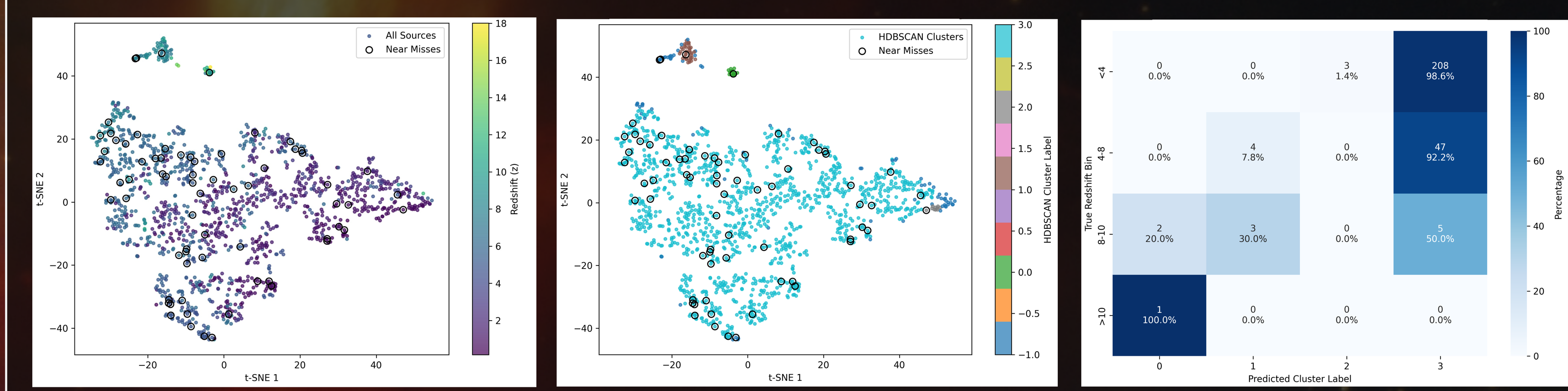


Figure 3. Architecture of our CNN, showing the flow from input multi-band images through spectral and spatial feature extraction. Spectral features are captured via 1×1 convolutions and SpectralResidualBlocks, refined by ChannelAttention. Spatial patterns are learned through SpatialResidualBlocks and MAXPOOLING, with final classification via global pooling and dense layers.

4. Semi-Supervised Approach & Results

Applying the t-SNE and HDBSCAN algorithms on magnitudes and flux radii isolate most $z \geq 8$ galaxies, with one cluster capturing nearly all $z \geq 10$ sources. A silhouette score of ~0.5 suggests moderate separation. A Random Forest trained on HDBSCAN labels classifies $z \geq 10$ well but overlap in the 4–10 range reveals limits of photometry for mid-redshift separation.



(4a) (4b) (4c)

Figure 5: (4a) t-SNE projection coloured by photometric redshift. (4b) HDBSCAN clusters overlaid. (4c) Confusion matrix from Random Forest trained on HDBSCAN labels. Bimodal galaxies (circled) are scattered.

3. Supervised Learning Results

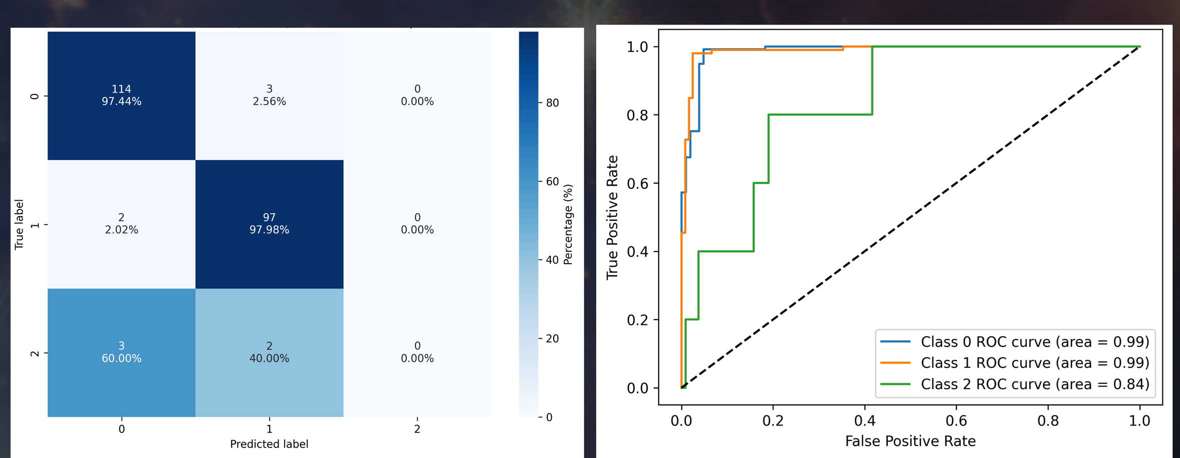
Framework: PyTorch
Input: 64×64×7-pixel cutouts
Optimiser: Adam, LR = 5e-4
Loss: Cross-entropy
Initialisation: He normal
Regularisation: Dropout, batch norm

The model rapidly converged within 10 epochs across 10 independent runs.

Typical results when using augmented training data and un-augmented test data:

- **Low-z accuracy**: ~97%
- **High-z accuracy**: ~98%
- **Brown dwarfs**: Poor separation: ~60% misclassified as low-z

Adding HOG features did not significantly improve performance, suggesting the CNN saturates its learning capacity from the original data.



(3a) (3b)

Figure 4. Confusion matrix (3a) and ROC curve (3b) from the best-performing CNN run. The model correctly classified 97.44% of low-redshift and 97.98% of high-redshift galaxies. All brown dwarfs were misclassified as galaxies.

Discussion

- **Ablation Study:** Removing photometric bands showed F200W is most critical (>45% accuracy drop), followed by F150W, F277W, and F356W (30–45%), highlighting the importance of spectral breaks. F444W has moderate impact; F115W and F090W contribute least.
- **Bimodal Redshifts:** 67 galaxies with bimodal PDFs (found using Balmer break aliasing criteria) challenge both pipelines. The CNN performs well (~99% TPR, ~96% TNR) but t-SNE/HDBSCAN fail to cluster them, revealing photometric limitations.
- **CNN Limitations:** The model separates spectral and spatial processing, limiting early spectral-spatial feature fusion. Fixed kernels, max-pooling, no spatial attention or positional encoding reduce sensitivity to complex features. No hyperparameter tuning was performed.
- **Brown Dwarfs:** Poorly classified due to class imbalance and few samples. Addressing this will require larger, more varied datasets and advanced techniques like few-shot learning or synthetic augmentation.
- **Misclassifications:** Low-redshift galaxies with bright centres and high background flux confuse the model, especially when spectral cues are weak. Using only the reddest band (F444W) drops accuracy to near chance, confirming reliance on spectral features like the Lyman break.

References

[1] V. Asadi, H. Haghi, et al., *Semi-supervised classification of stars, galaxies and quasars using k-means and random forest*, Astronomy & Astrophysics, 2025.

[2] F. D'Eugenio, A. J. Cameron, et al. *Jades data release 3 – NIRSpec/MSA spectroscopy for 4,000 galaxies in the GOODS fields*, The Astrophysical Journal, 2024.

[3] D. Austin et al., *Galaxy catalogue for the upcoming EPOCHS v2 paper,* In Prep, https://github.com/duncanaustin98/galfind.git.

[4] D. Austin, *Inferring the Properties of Star-Forming Galaxies in the Epoch of Reionization with JWST*, PhD thesis, University of Manchester, In Prep.

[5] T. Harvey and D. Austin, *Brown dwarf fitter GitHub repository*, 2025, https://github.com/tHarvey303/BD-Finder.git.