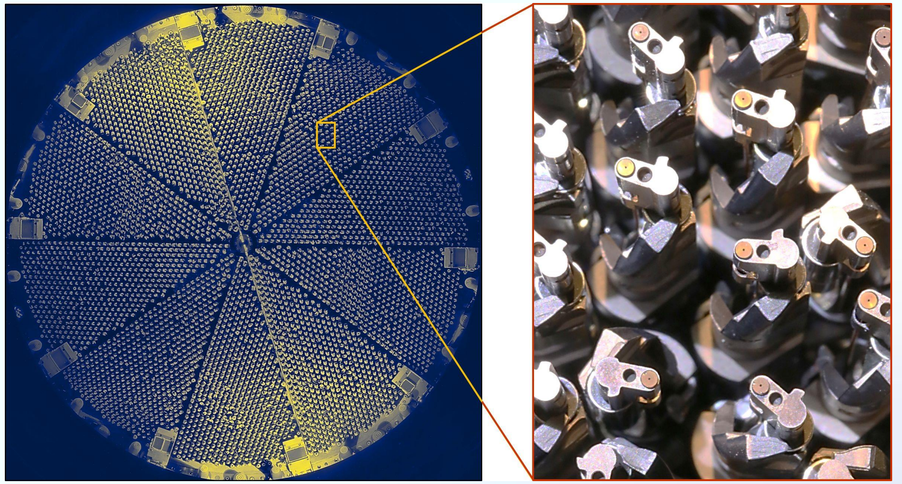
AstroVision Model for Imaging and Spectroscopy (AVIS)

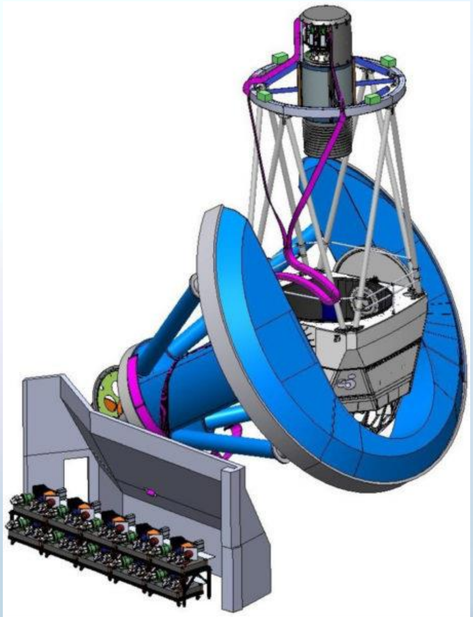
Yufeng Luo
Advisors:
Adam Myers, Stéphanie Juneau
July 15, 2026
2nd Cosmic-Horizons Conference

Dark Energy Spectroscopic Instrument

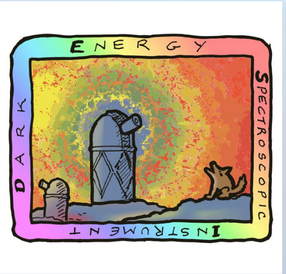
- The Dark Energy Spectroscopic Instrument (DESI) is mounted on Mayall 4-meter telescope at Kitt Peak National Observatory
  - Has 5000 robotically controlled fibers for spectroscopy
  - Targeting various tracers for cosmology
  - Wavelength range from 3600Å to 9800Å with blue, red, and near-infrared spectrographs
  - Data Release 1 includes deep spectra for 18.7 million unique sources (extragalactic and MW)
- Targets are selected via a large area sky survey -- DESI Legacy Imaging Surveys (LS)



Top: DESI Focal plane and a close view of the robotic fibers. Credit: https://www.desi.lbl.gov/for-scientists/instrument-design

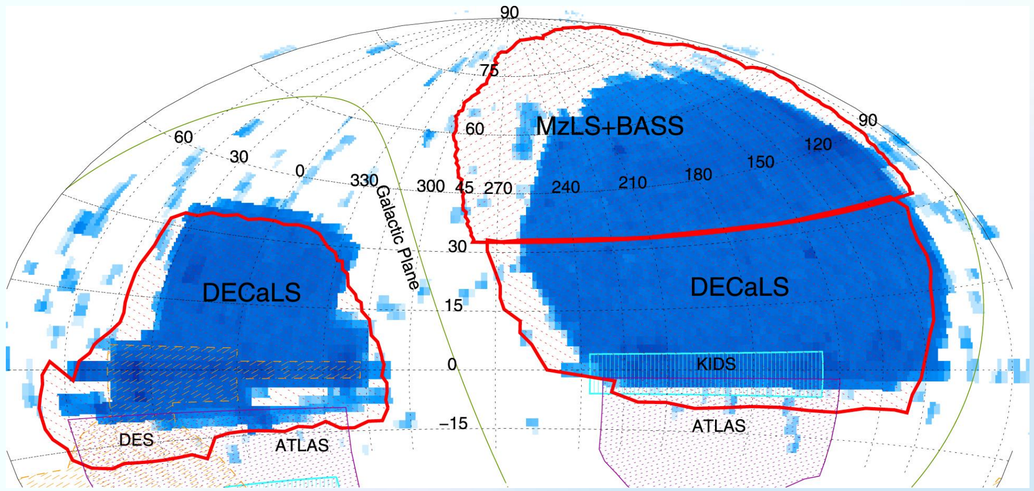


Left: DESI design on the Mayall telescope. Credit: https://www.desi.lbl.gov/instrument/



DESI Legacy Imaging Surveys

- LS consists of three surveys: Beijing-Arizona Sky Survey (BASS), Mayall z-band Legacy Survey (MzLS), and DECam Legacy Survey (DECaLS)
- Covers ~14,000 deg^2 sky area in grz bands
- The number of unique sources in Data Release (DR) 10 is 2×10^9
- DR9 is used for DESI main survey target selection
- DR11 will incorporate new data from DECaLS for DESI-II target selection, and is being processed
  - Need to identify problematic exposures



The footprint of LS as bounded by the three red curves. Credit: Dey+2019.

DESI and Legacy Surveys (LS)

- Both surveys contain a large amount of data for diverse sources
- Each source in DESI catalog can always be matched in LS catalog
- **Challenges** in DESI and LS
  - Data quality: Bad exposures can be difficult to identify in a large survey like LS (Luo+ 2025)
  - Computation: Analyzing the catalogs can be computationally challenging using traditional methods.
  - Labels: the number of labels can be limited and might be biased, depending on the predefined templates
- **Opportunity**: The two datasets provide an unique platform to develop and test innovative algorithms, such as advanced deep learning
  - Fast source classification and estimation on various physical properties, such as morphology and redshift in a large observational campaign (DESI, LSST, etc)
  - Discovery of hidden connections between physical properties
  - Extensible framework to enable a new approach to problems

The Rise of Foundation Models

- Foundation models (FM) are trained via self-supervised learning (SSL) on massive dataset and can be adapted for various downstream tasks after fine-tuning.
  - SSL does not need human-expert labels
  - Learning the representation of the data in high-dimensional space
- Large-Language Models (LLM), such as ChatGPT, are good examples
- Multi-modal foundation models can combine multiple information together and potentially gain deeper understanding of the correlation between modalities

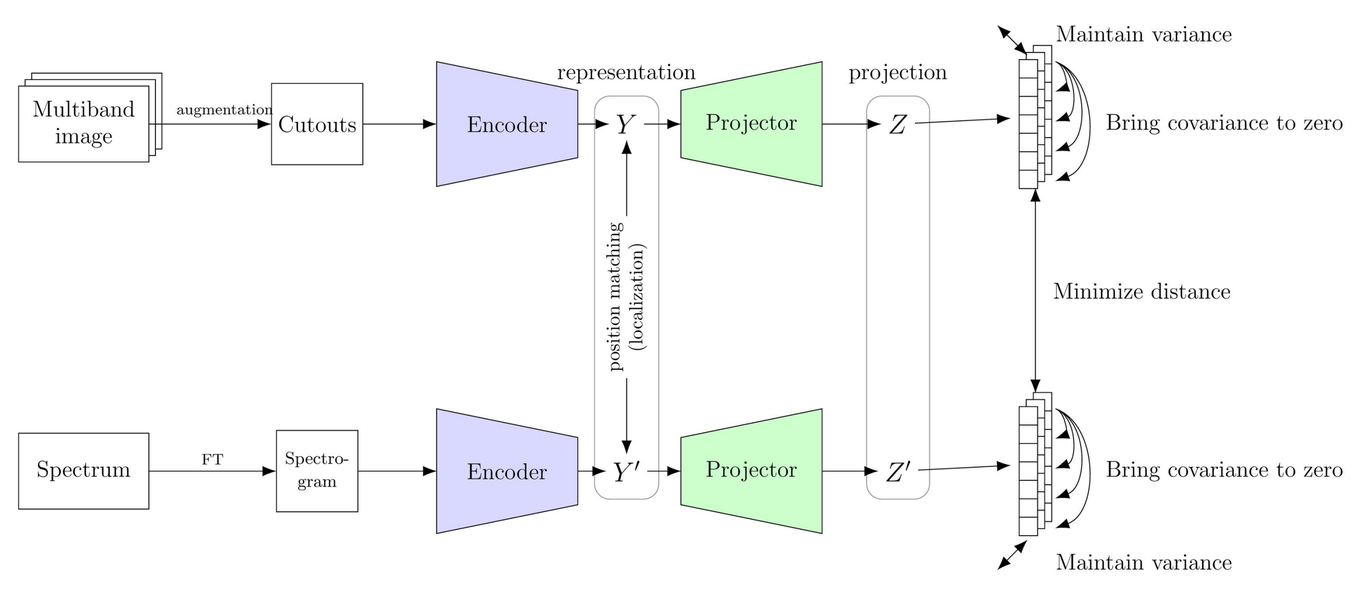
ChatGPT

Claude

Gemini

AVIS: AstroVision Model for Imaging and Spectroscopy

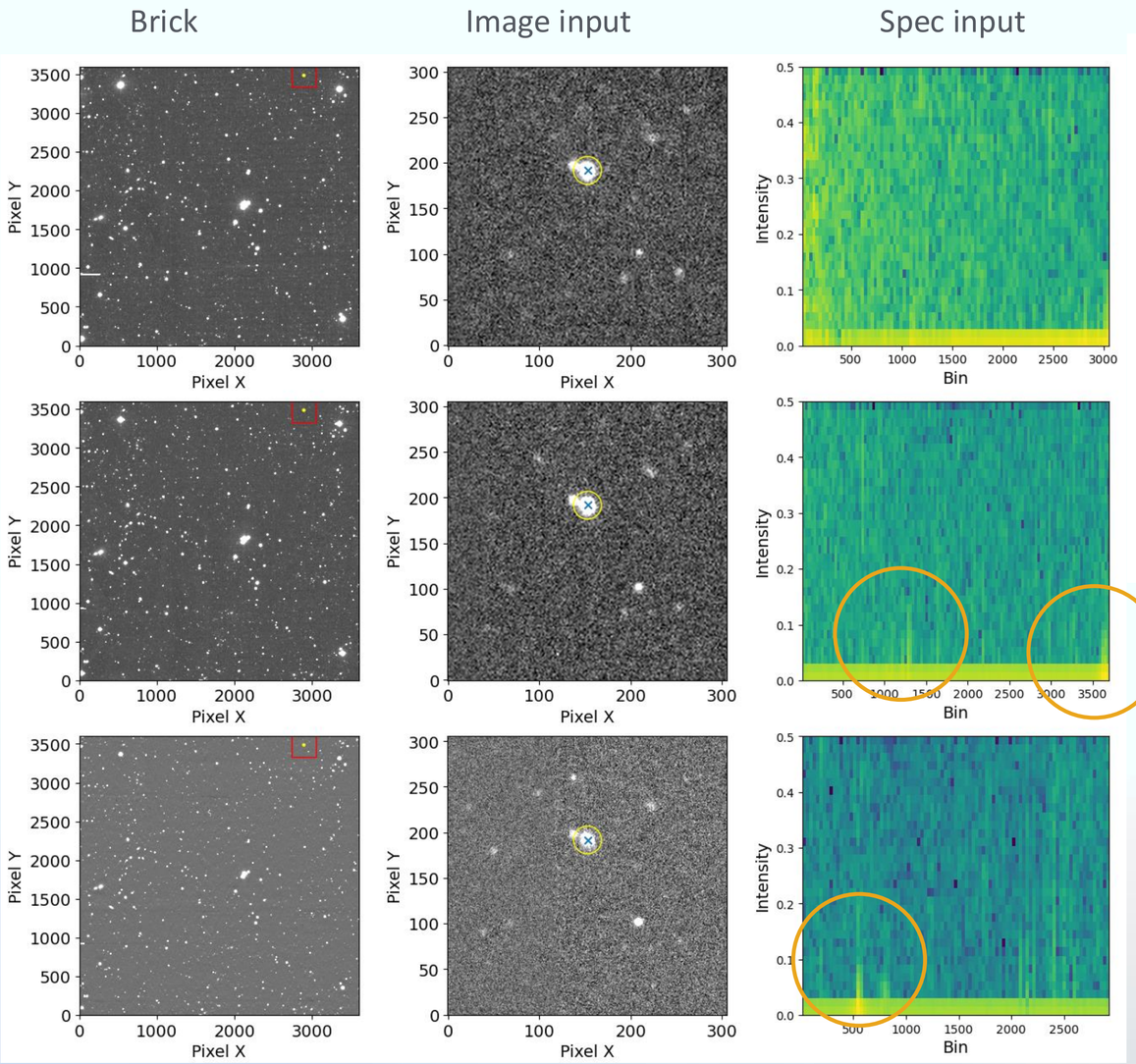
- Based on Variance Invariance and Covariance Regularization algorithm (VICReg; Bardes+2022) with modifications on image and spectrum pair input. **No human expert** label needed for training, **non-contrastive**
- Additional loss to help model learn correlational features
- Keep the same deep learning model (ViT) for both image and spectrum for simplicity of verification
- Trained on 3 million sources subsampled from all DESI Y3 data

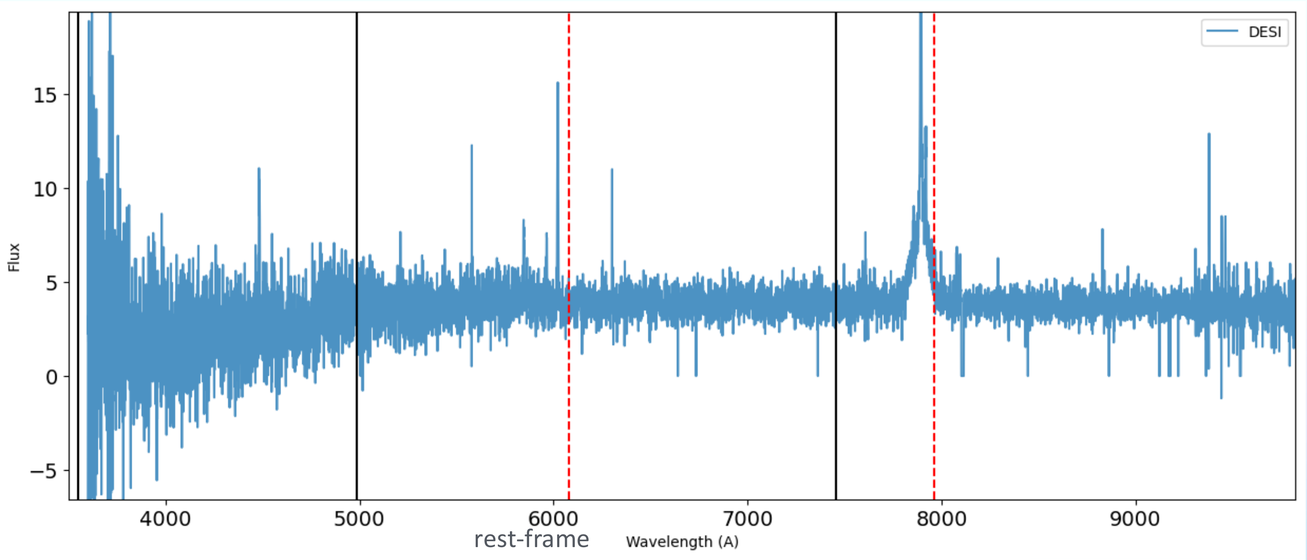


$$f(image, spec) \rightarrow z \in \mathbb{R}^n$$

$$\mathcal{L} = \mathcal{L}_{VICReg} + \mathcal{L}_{DESI}$$

Data Preparation





A DESI Quasar with Broad Hα line

Left: grz band images and spectrogram of a source , x-axis is bin number and y-axis is the intensity

top: DESI spectrum of the source. Black and red lines corresponding to the bandwidth of the imaging filters

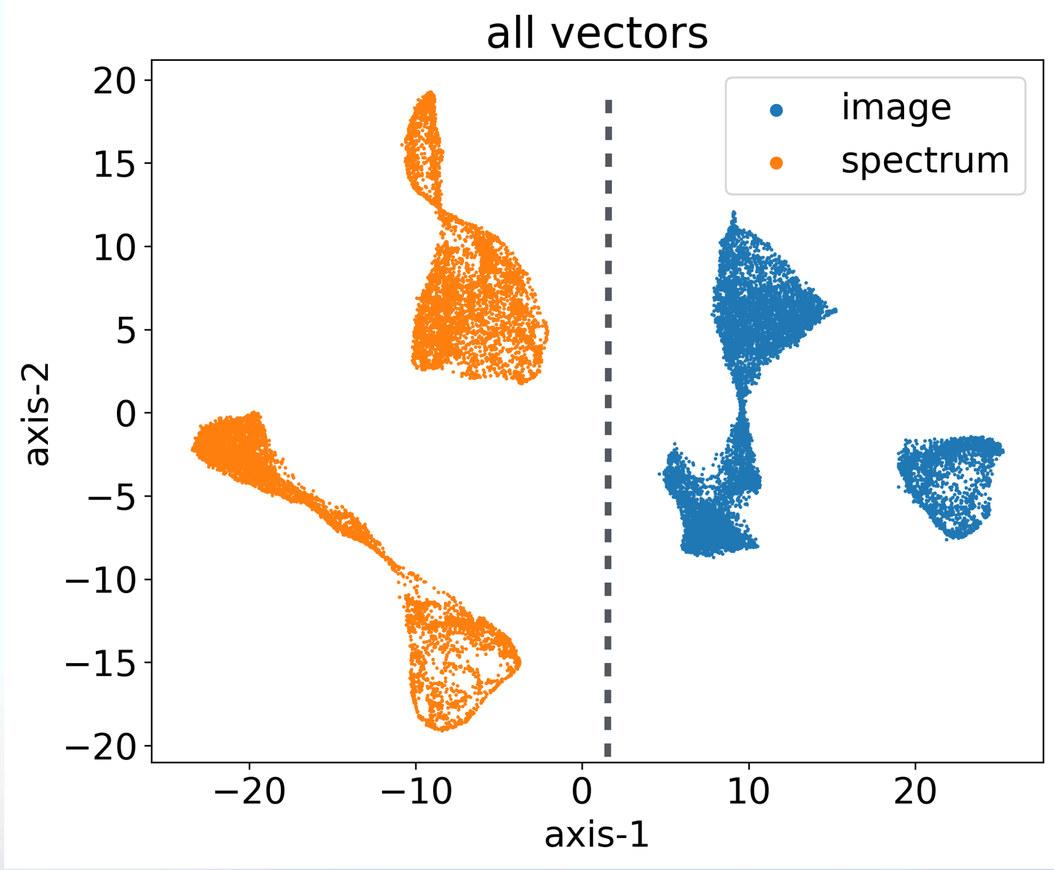
Analysis of Embedding Space

- Purpose: **verify** the model has learned some physical insights from the dataset and provide a preliminary **understanding** of meanings of the clusters
- Original outputs from the model are embedding vectors with 384 dimensions.and are reduced to 2 via Principal Component Analysis (PCA) and Pairwise Controlled Manifold Approximation (**PaCMAP**) for visualization and analysis
- What to look for:

$$f_{SSL}(image, spec) \rightarrow z \in \mathbb{R}^{384}$$

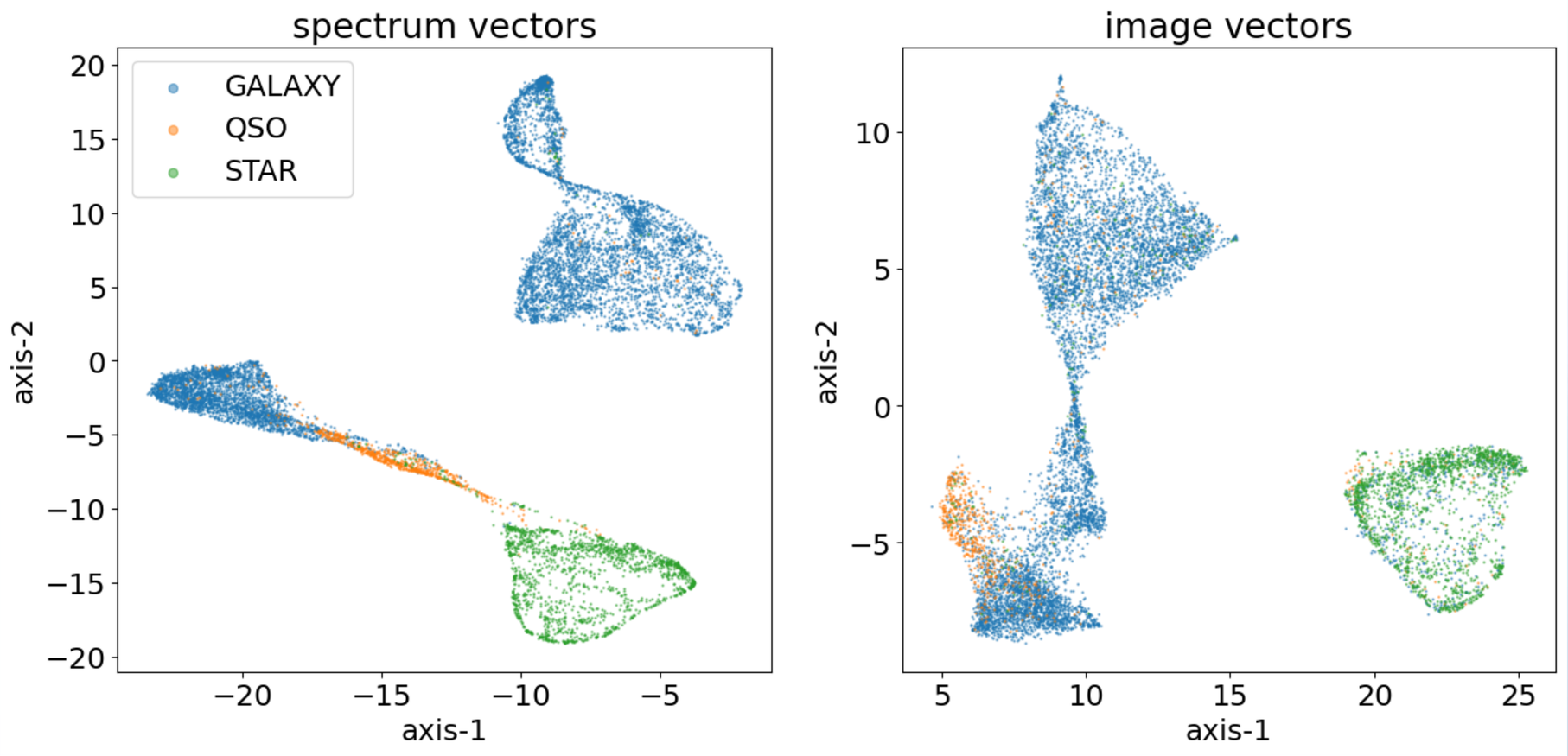
$$g_{PM}(z) \rightarrow z' \in \mathbb{R}^{2}$$

  - Axes do not have extrinsic physical meanings, they are a combination of various latent features
  - **Separations**: If clusters are clearly separated, it means model can distinguish different sources
  - The **geometries** and **relative locations** of clusters usually show connections between characteristics of various categories of sources



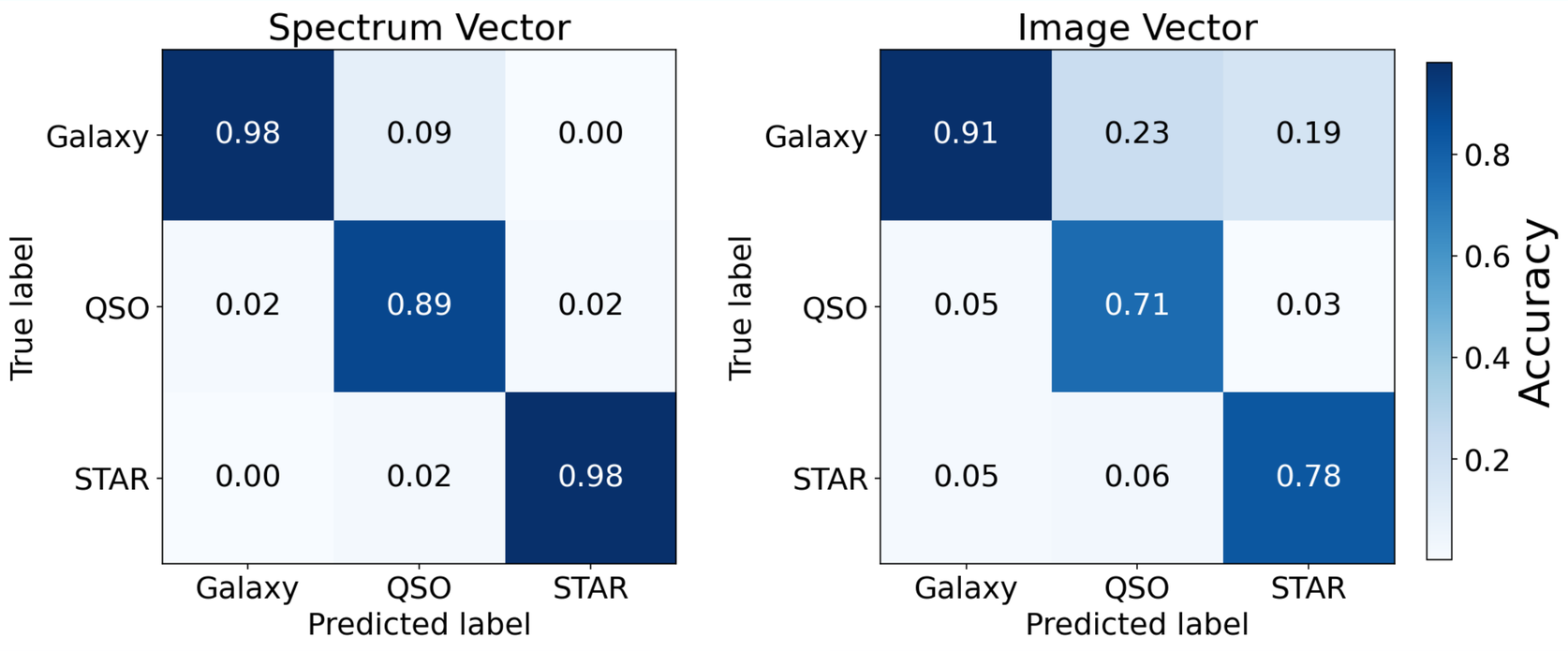
Spectral Template Type

DESI has three spectral type for the spectra-fitting pipeline: Galaxy, QSO and Star



Classification on spectral type

- k-Nearest Neighbor (kNN) Classifier with imaging and spectroscopic embeddings together
- Good out-of-the-box performance before any fine-tunings or continual learning



Definition of DESI Targets

- DESI's observation targets include:
  - Four main extragalactic tracers for cosmology:
    - **Bright Galaxies** (Bright Galaxy Survey, BGS) -- nearby bright galaxies (r < 20.175 mag)
    - **Luminous Red Galaxies** (LRG) -- giant elliptical galaxies with older stellar population and no active star formation
    - **Emission Line Galaxies** (ELG) -- strong emission lines (e.g., [OII]) from active star formation
    - **Quasars** (QSO) -- accreting SMBH
  - Milky Way Sources (Milky Way Survey, MWS)
  - Other secondary targets (SCND), can be various kinds, such as Lyman Alpha Emitter (LAE), Lyman Break Galaxies (LBG) and transients, etc.

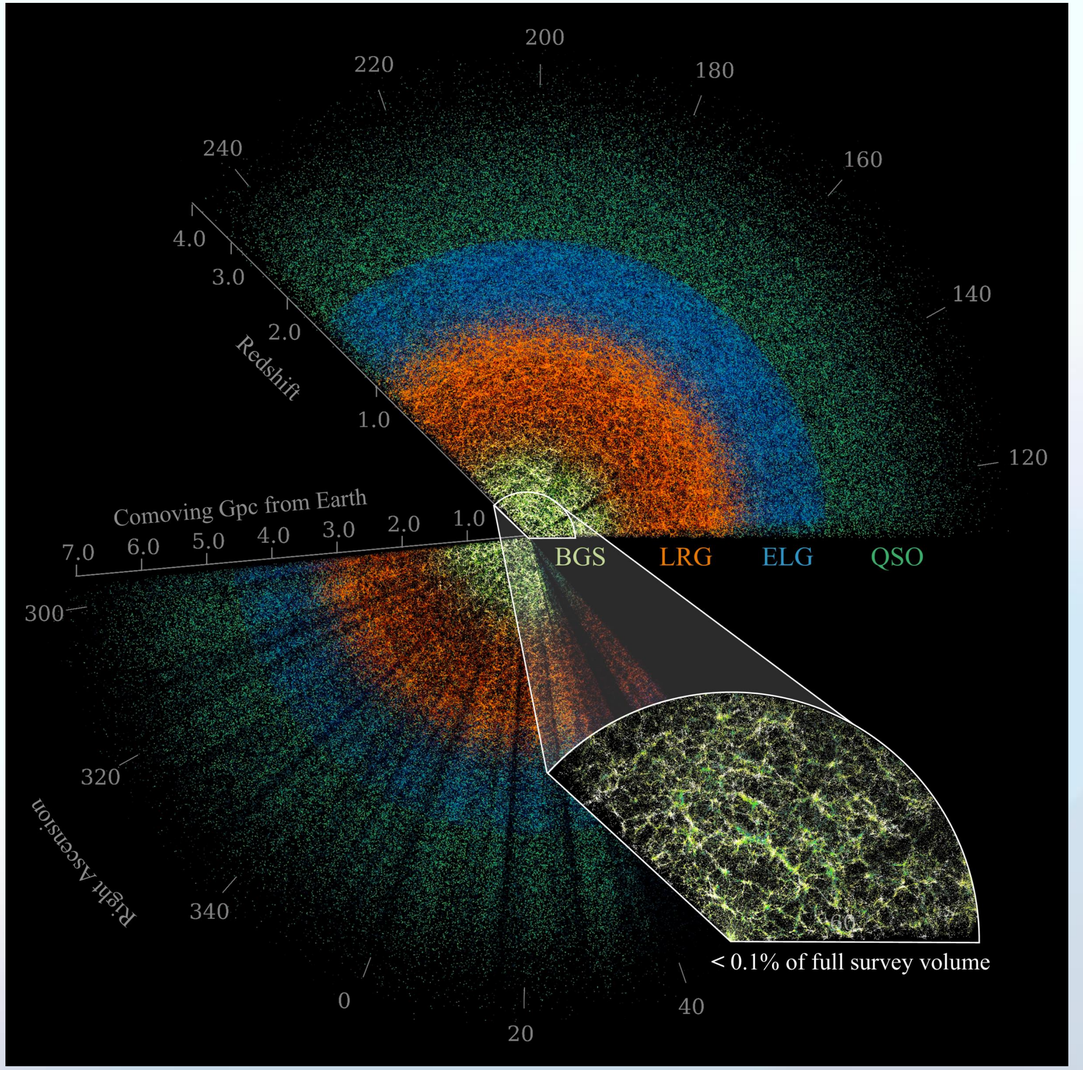
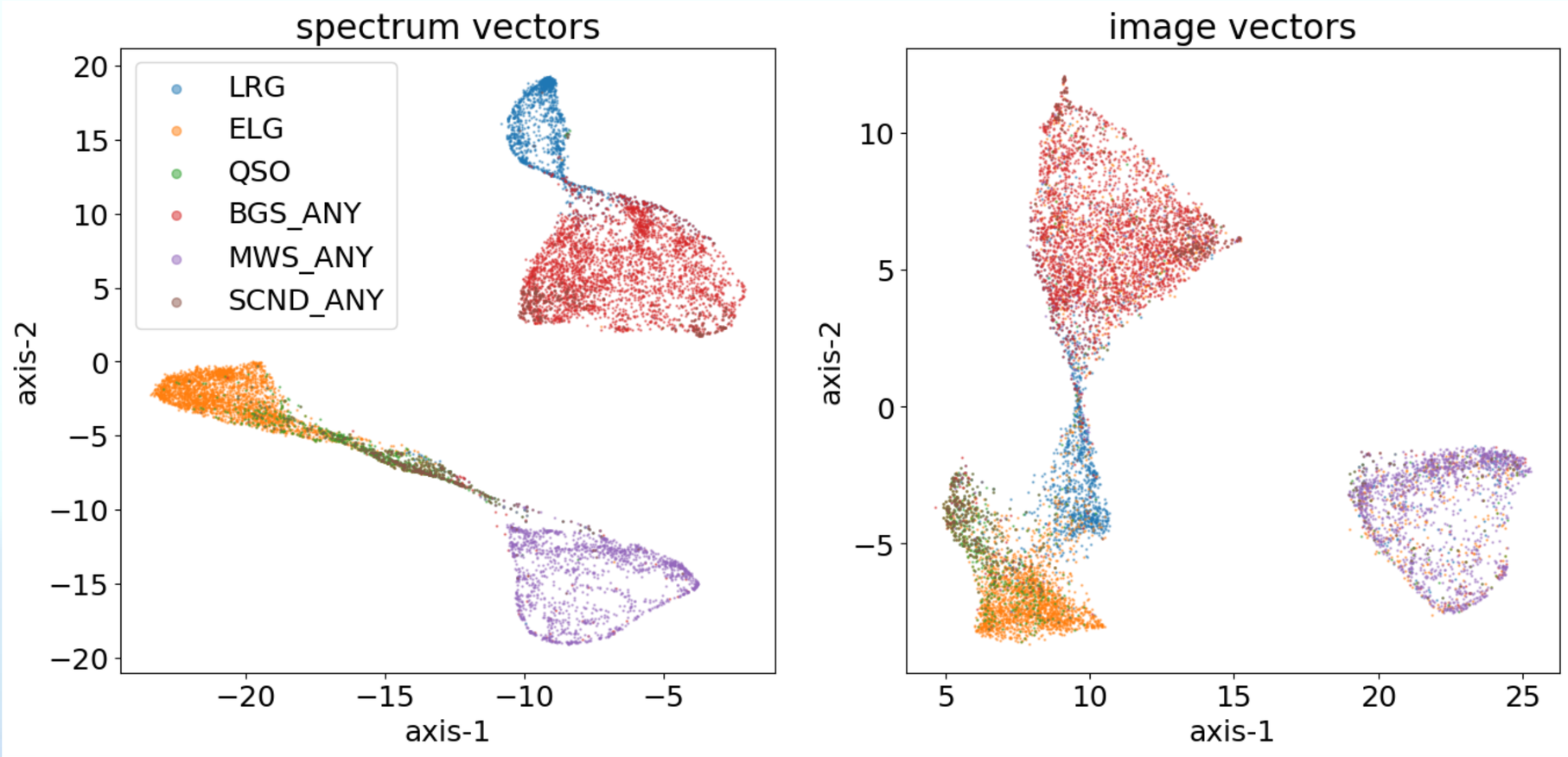


Fig 1. of DESI DR1 paper, showing the distribution of tracers.
Credit: DESI Collaboration+2025

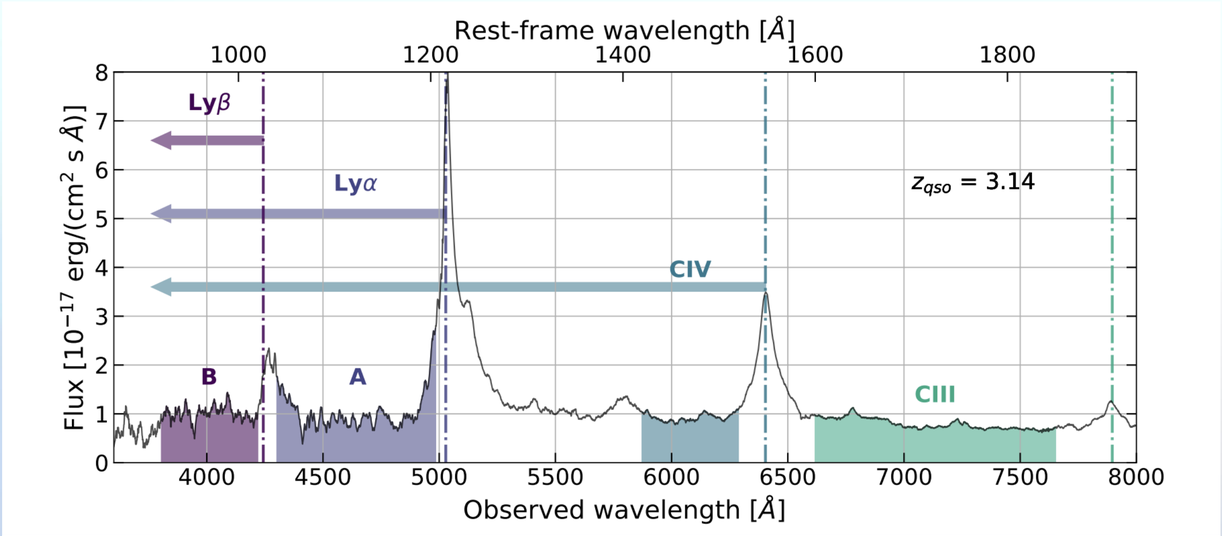
DESI Target in Embedding Space

Clear separation between each clusters

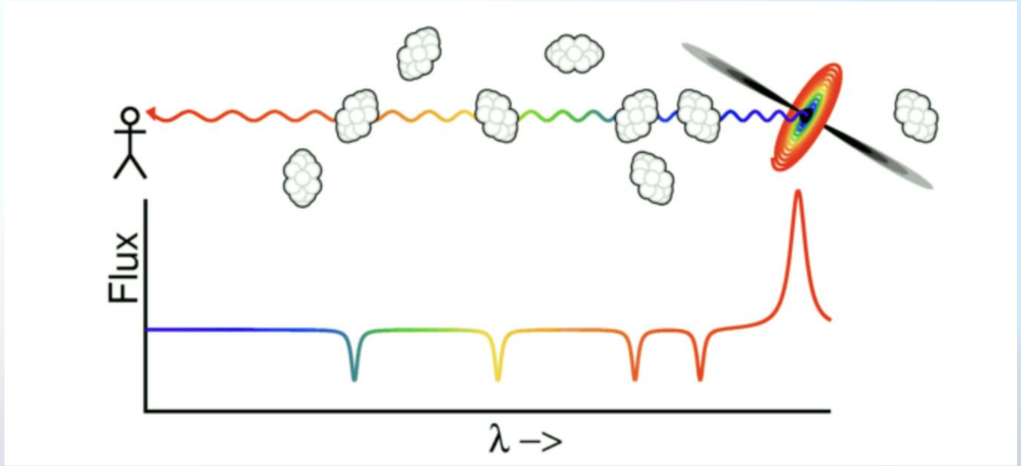


Science Application: LyA Forest QSO

- Lyman Alpha (LyA, $\lambda 1215\ \text{Å}$) is one of the most prominent emissions to probe high-redshift galaxies and IGM
- Emission: Background QSO emits LyA photons
- Absorption: Intervening gas clouds along line-of-sight as LyA shifts toward redder end
- Lyman Alpha Forest QSO can be used to measure the Baryon Acoustic Oscillation (BAO) signal



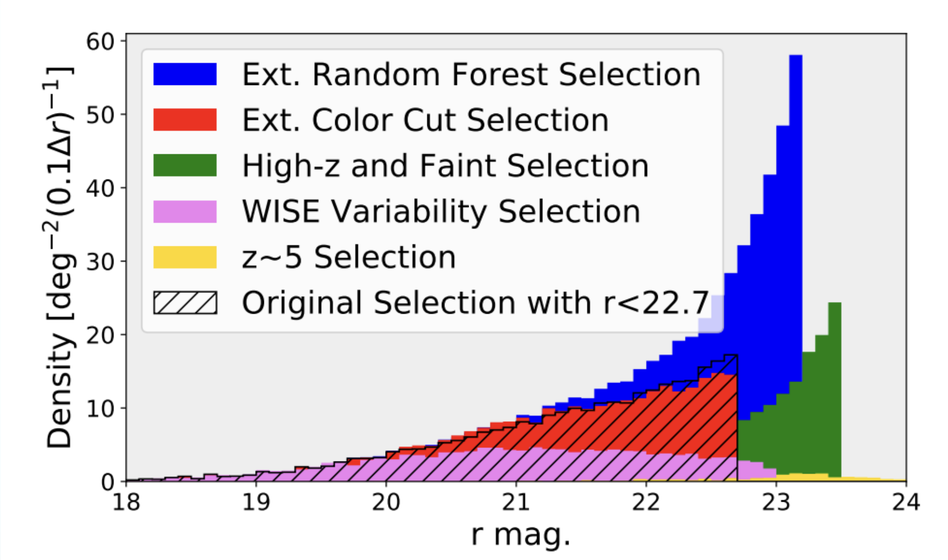
DESI Collaboration+2024: BAO from Lya forest



Credit: https://www.astro.ucla.edu/~wright/Lyman-alpha-forest.html

Target Selection and Spectra Classification

- LyA forest QSO are selected as a subclass to general QSO target class

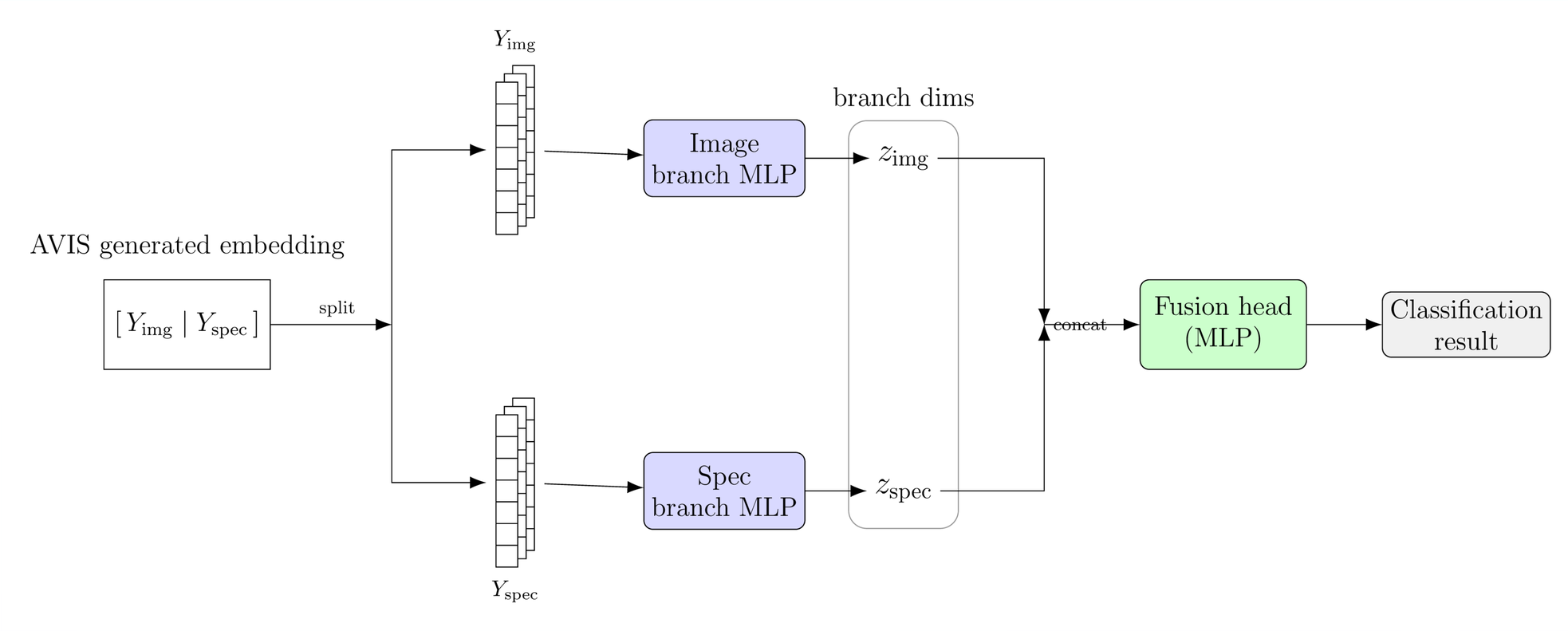


Chaussidon+ 2022

- Hard for template fitting codes because of the absorption features on the blue continuum
- There are other features such as Broad-Absorption Line (BAL) that need special treatment
- DESI's CNN-based QuasarNet has a very good performance (purity/precision ~99.5%) for all QSO targets, but it only uses spectra

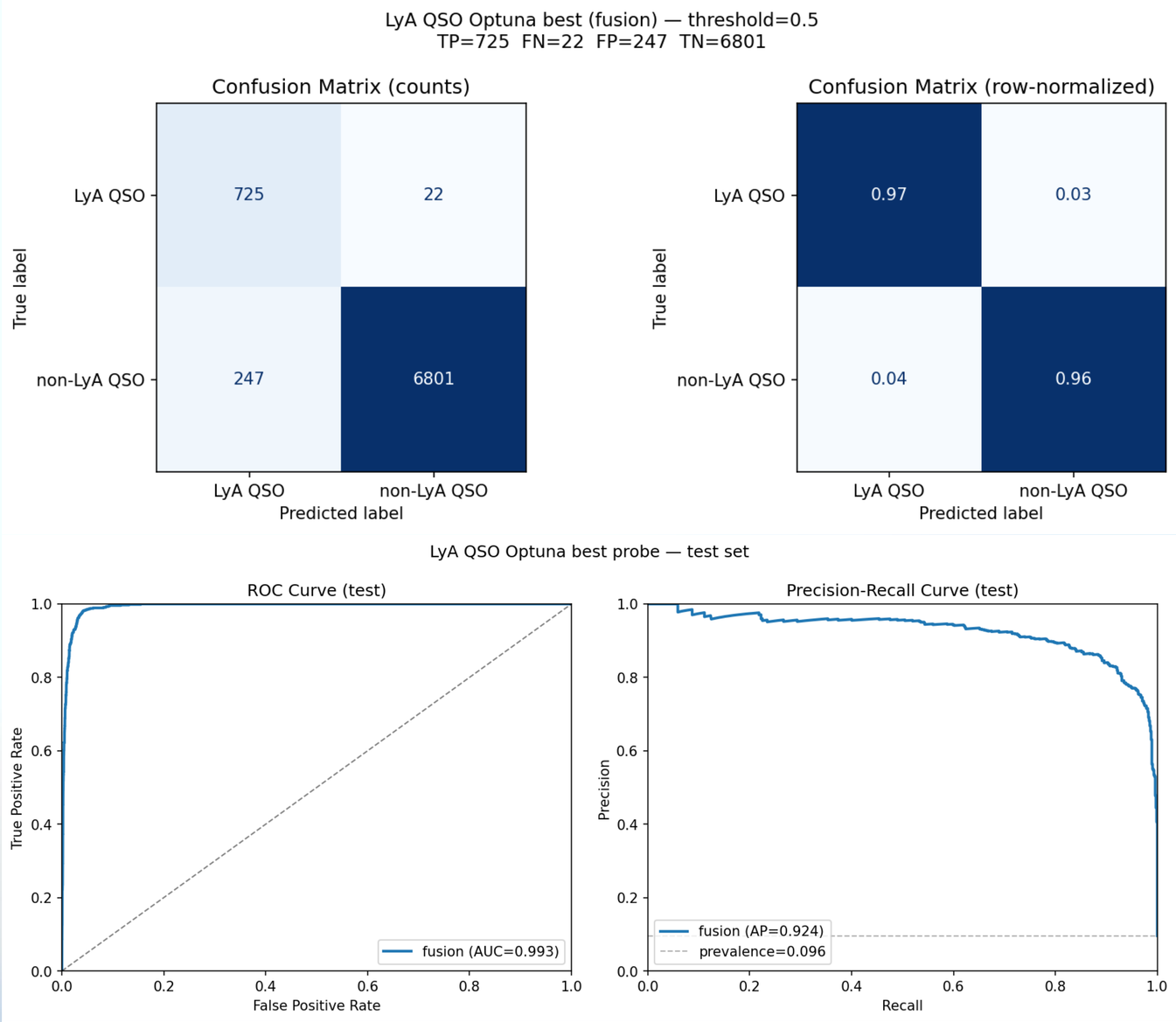
Architecture

- Supervised downstream science task using embeddings generated via pre-trained AVIS
- Two branch + fusion MLPs



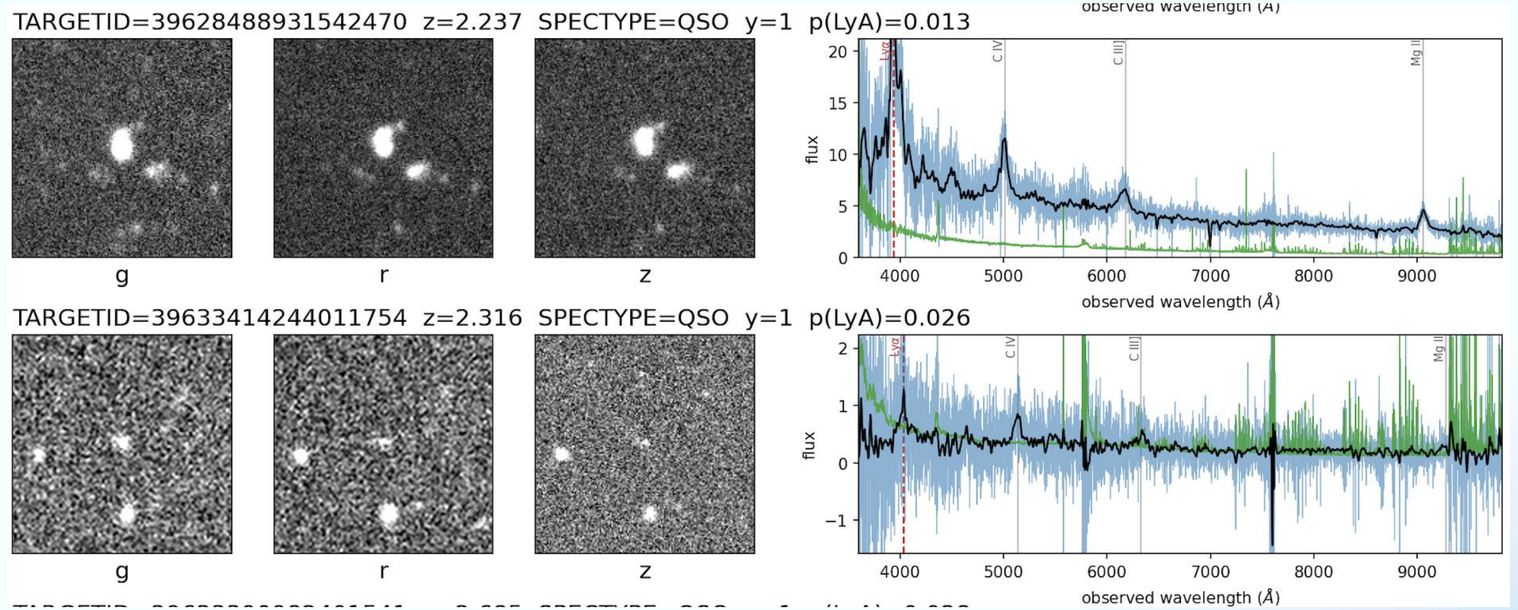
Classification Result

- Trained on 10k sources
  - 10% pure Lya-forest QSO sources
  - Negative includes (BALs,) low-z QSO, galaxies, stars
- Highly-imbalanced classification
- Average Precision at 0.924 for Precision Recall curve

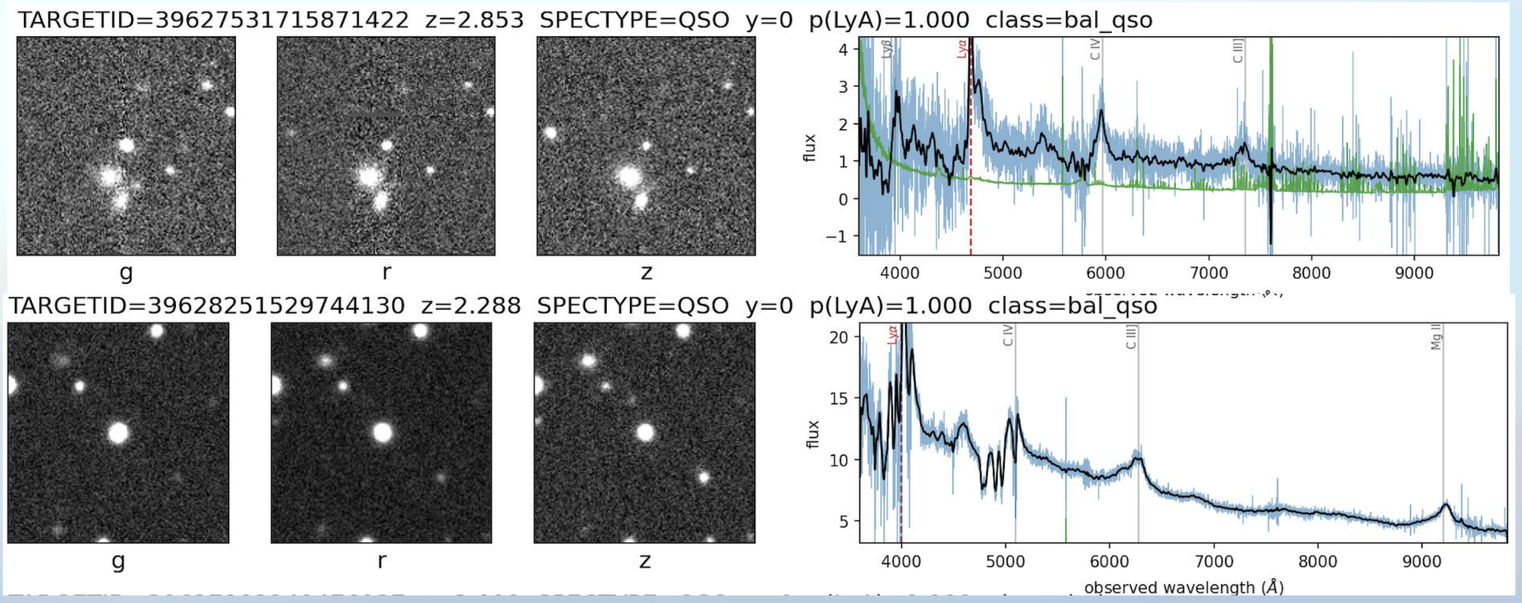


Some cases from FP and FN samples

False Negatives



False Positives



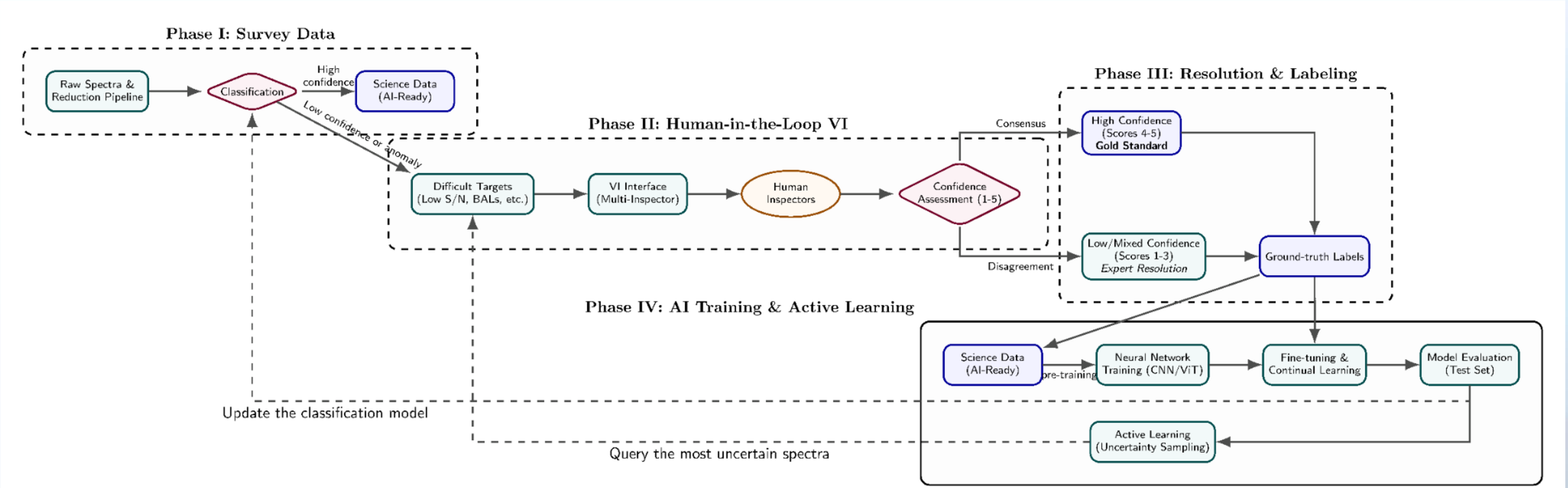
Conclusion

- Demonstrated the application of computer vision-based foundation model to imaging and spectroscopic data via self-supervised learning
- By clustering in the embedding space, the model displays capabilities to connect different physical properties.
- The classification performance further confirms the model has ability to distinguish sources using these connections
- Achieved good performance on downstream tasks, such as Lya forest QSO classification task
- Further studies are needed to understand the full capabilities of the model and knowledge it has learn. We need metric and benchmark to do this
- Future opportunities: other downstream science tasks. Transfer/continual learning to different data set such as LSST and Roman data

Applying for postdoctoral positions this fall!

One More Thing....

- DESI is entering DESI-1b and preparing for DESI-II
- Visual Inspection becomes essential with new target classes (e.g., LAEs) and anomalies
- How can we use ML models to assist in the VI process?



Thank you for listening!

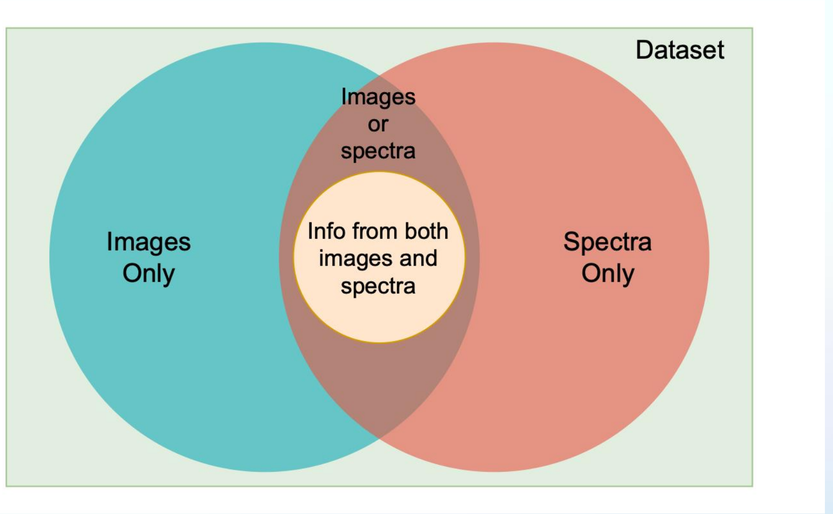
Backup Slides

Considerations for FM in Astronomy

- **Data modality**: what data should we use?
- **Training philosophy**:
  - Contrastive learning: finding the similarities between positive pairs while pushing negative pairs away
  - Non-contrastive learning: only focus on similarities between modalities
- **Evaluation**: How can we interpret and explain the model and ensure trustworthiness of the model within astrophysical context?
- **Application**: How to adapt the model to solve astrophysical problems?

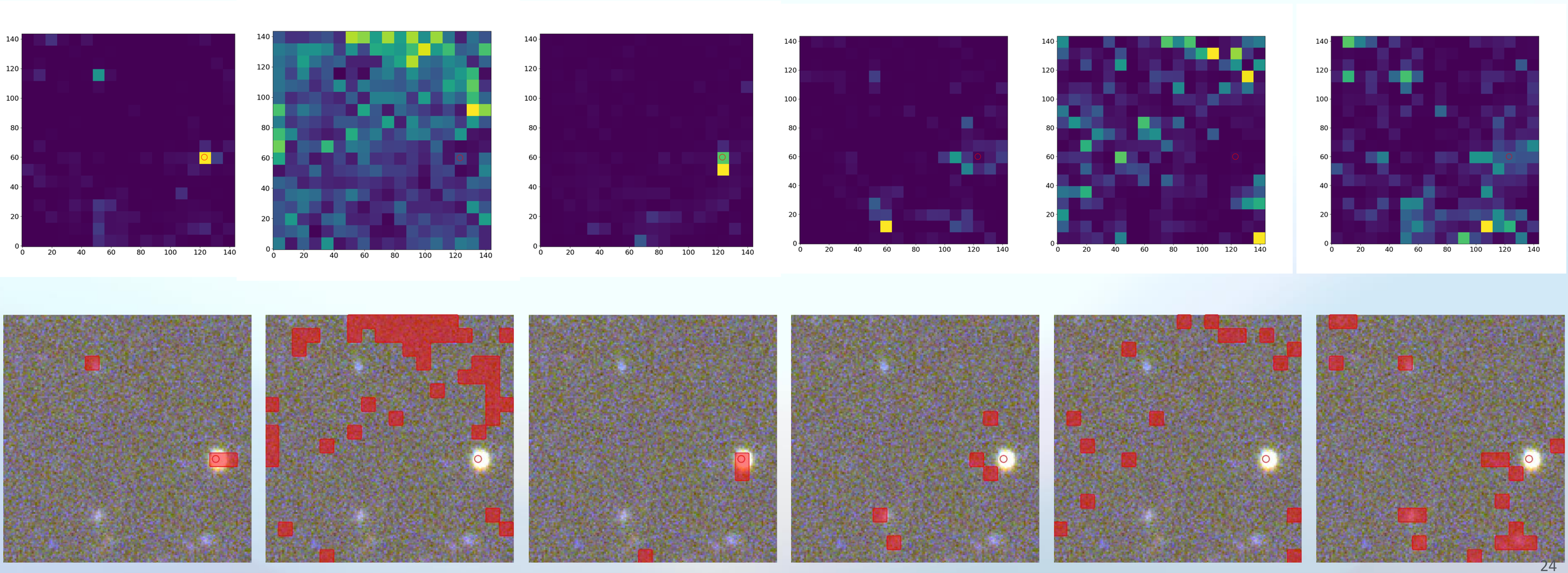
Evaluation Consideration: Benchmarking

- Need to understand general performance of the model in a **standard**, **interpretable** way
- Solving some (simple) astrophysical tasks with a standard multimodal dataset
  - Tasks that can be accomplished using only imaging data.
  - Tasks that can be accomplished using only spectra.
  - Tasks that can benefit from both imaging and spectra if they are available.
  - Tasks that require both imaging and spectroscopic data.
- As a sample characterization TG lead, I plan to host several benchmark challenges, similar to the stellar mass challenge / black hole mass challenge
- Visualizing the high-dimensional embedding vectors and correlation with other data
  - can we use glue?



Attention Map: Images

- QSO TARGETID: 39628523215786289



General Workflow

- Loading and preprocessing
  - Full brick-based images, do cutouts on-the-fly
  - Coadded spectra are transformed to spectrograms on the fly
- Spectra are in HDF5 and images are in FITS
- Embeddings are in HDF5