

Slitless Spectroscopy Source Detection Using YOLO Deep Neural Network

XIAOHAN CHEN ^{1,2} MAN I LAM ^{*,†,1} YINGYING ZHOU ³ HONGRUI GU ^{1,4} JINZHI LAI ^{1,4} ZHOU FAN ^{*,1,4} 
JING LI ^{*,2}  XIN ZHANG,¹ AND HAO TIAN ¹

¹*National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, People's Republic of China*

²*School of Physics and Astronomy, China West Normal University, Nanchong 637009, People's Republic of China*

³*Key Laboratory of Space Astronomy and Technology, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, P.R. China*

⁴*School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China*

ABSTRACT

Slitless spectroscopy eliminates the need for slits, allowing light to pass directly through a prism or grism to generate a spectral dispersion image that encompasses all celestial objects within a specified area. This technique enables highly efficient spectral acquisition. However, when processing CSST slitless spectroscopy data, the unique design of its focal plane introduces a challenge: photometric and slitless spectroscopic images do not have a one-to-one correspondence. As a result, it becomes essential to first identify and count the sources in the slitless spectroscopic images before extracting spectra. To address this challenge, we employed the You Only Look Once (YOLO) object detection algorithm to develop a model for detecting targets in slitless spectroscopy images. This model was trained on 1,560 simulated CSST slitless spectroscopic images. These simulations were generated from the CSST Cycle 6 and Cycle 9 main survey data products, representing the Galactic and nearby galaxy regions and the high galactic latitude regions, respectively. On the validation set, the model achieved a precision of 88.6% and recall of 90.4% for spectral lines, and 87.0% and 80.8% for zeroth-order images. In testing, it maintained a detection rate $>80\%$ for targets brighter than 21 mag (medium-density regions) and 20 mag (low-density regions) in the Galactic and nearby galaxies regions, and $>70\%$ for targets brighter than 18 mag in high galactic latitude regions.

Keywords: Astronomical object identification (87) — Deep learning (1938) — Spectroscopy (1558)

1. INTRODUCTION

Slitless spectroscopy, produced by a prism or grism without the use of a slit, is an efficient survey tool because it records the spectrum of all objects within the field of view. It is also more straightforward than traditional methods, as it eliminates the need for pre-imaging or slit alignment. Observations can be conducted simply by positioning the target within the field of view, greatly simplifying the observational process and reducing the need for precise telescope pointing. Additionally, traditional slit spectroscopy can result in light loss due to parts of the object not passing through the slit, which reduces the signal-to-noise ratio (SNR) of the spectrum. In contrast, slitless spectroscopy can minimize light loss by allowing more photon to reach the detector, resulting in more accurate flux calibration (C. J. Willott et al. 2022).

Due to its operational simplicity and high efficiency in spectral acquisition, slitless spectroscopy has become a widely used technique in both ground-based and space-based observational astronomy. In ground-based surveys, R. G. Clowes et al. (1980) employed objective prism plates to perform low-resolution spectroscopic measurements of quasars, a technique that has been crucial for studying the statistical properties, redshift distribution, and spectral characteristics of quasars. This method allows researchers to efficiently obtain spectral data for large numbers of quasars without the need for massive telescopes. The efficiency of this technique paved the way for several major quasar surveys. Earlier, wide-field, objective prism-based slitless spectroscopic surveys, such as the Hamburg Quasar

Survey (HQS) (D. Engels et al. 1988; H. J. Hagen et al. 1995; D. Engels et al. 1998) and the Hamburg/ESO Survey (HES) (L. Wisotzki 1994; L. Wisotzki et al. 1996; D. Reimers et al. 1996; D. Reimers & L. Wisotzki 1997), discovered a large number of bright quasars. Later, G. Worseck et al. (2008) employed a deeper, targeted, CCD-based slitless spectroscopic survey — the ‘Quasars near Quasars’ (QNN) survey — to discover a significant number of fainter quasars at intermediate and high redshifts.

Slitless spectroscopy is particularly efficient in space-based surveys due to the reduced background noise in space (M. Kümmel et al. 2009). Space telescopes such as the Hubble Space Telescope (HST) (A. Pasquali et al. 2006), the James Webb Space Telescope (JWST) (T. P. Greene et al. 2016; C. J. Willott et al. 2022; R. Doyon et al. 2023), and Euclid (Euclid Collaboration et al. 2025a) are all equipped with slitless spectroscopic modules. Leveraging the near-infrared slitless spectrographs (G102 and G141) on its Wide Field Camera 3 (WFC3), the HST observes the central regions of galaxy clusters to obtain spectral data for high-redshift galaxies. These observations advance studies in galaxy formation and evolution, star formation history, chemical composition, and the early universe. For example, observations of the Spiderweb protocluster core using the G141 grism on HST WFC3 revealed the important role of active galactic nuclei (AGN) in influencing the evolution of quiescent galaxies (A. Naufal et al. 2024). Moreover, by using HST WFC3/IR imaging and spectroscopic data, researchers successfully identified galaxies at redshifts as high as $z \approx 11.09$, providing key constraints on galaxy assembly processes in the early universe (P. A. Oesch et al. 2016). Another study combined deep slitless spectroscopic and imaging data from multiple HST surveys to investigate the star formation history and metallicity evolution of massive galaxies that had already quenched when the universe was only ~ 2 Gyr (T. Morishita et al. 2019). Similarly, the Wide Field Slitless Spectroscopy (WFSS) mode of the Near Infrared Imager and Slitless Spectrograph (NIRISS), together with the Near Infrared Camera (NIRCam) on JWST, enables slitless spectroscopic observations across the $0.8\text{--}2.2\ \mu\text{m}$ and $2.4\text{--}5.0\ \mu\text{m}$ wavelength ranges, respectively. With its powerful capability, JWST/NIRCam can observe high-redshift galaxies to study their metallicity, ionizing photon production efficiency, physical conditions of the interstellar medium, and the evolutionary characteristics of galaxy luminosity functions in the early universe, thereby revealing the formation and evolution mechanisms of galaxies during the Epoch of Reionization (F. Sun et al. 2023). And JWST/NIRISS achieves a depth and spatial resolution higher than ever before, which is particularly crucial for resolving individual stars or galaxies in crowded cluster environments (K. Boyett et al. 2022). Furthermore, JWST/NIRISS is capable of providing high-quality data for research in the field of exoplanets, assisting scientists in interpreting the atmospheric composition and physical conditions of exoplanets, as well as inferring their origins and evolutionary pathways (C. Piaulet-Ghorayeb et al. 2024). Following in the footsteps of HST and JWST, Euclid is also set to contribute significantly to our understanding of the universe. Using data from the Euclid Quick Data Release (Q1) (Euclid Collaboration et al. 2025b), the Near-Infrared Spectrometer and Photometer (NISF) aboard the Euclid satellite has enabled significant findings across numerous fields such as nearby galaxies, galaxy morphology, and strong gravitational lensing. A notable demonstration of NISF’s capability is the discovery of a high-redshift quasar ($z = 5.404$), EUCL J181530.01+652054.0, using its slitless spectrograph ($R \gtrsim 450$ across $950\text{--}2020\ \text{nm}$) (Euclid Collaboration et al. 2025c; Y. Copin et al. 2025; E. Bañados et al. 2025).

Although slitless spectroscopy significantly enhances observational efficiency by eliminating the need for a slit, it also introduces contamination effects that complicate spectrum extraction. Currently, there are several specialized data processing pipelines designed for slitless spectroscopy. For example, GRIZLI (G. Brammer 2019) focuses on comprehensive modeling and fitting of slitless spectroscopic observations; LINEAR (R. E. Ryan et al. 2018b,a) reconstructs high SNR one-dimensional spectra by solving linear equations while effectively handling contamination issues; and aXe (M. Kümmel et al. 2009) is a versatile slitless spectroscopy data processing tool specifically developed for the HST, supporting multiple instruments and featuring advanced spectral extraction and correction capabilities. Contamination is a common challenge in slitless spectroscopy. In cases with a high source density, contamination effects become more frequent and severe, making spectral extraction more difficult. In such scenarios, the scientific value of the data can be compromised. Therefore, it is crucial to detect and count the sources in slitless spectroscopic images, excluding data with excessively high source densities (i.e., data with significant contamination) before proceeding with spectral extraction.

The China Space Station Telescope (CSST) is equipped with a slitless spectroscopy module and is set to conduct a large-scale slitless spectroscopic survey. This survey will cover approximately $17,500$ square degrees of celestial area with a wavelength range of $255\text{--}1000\ \text{nm}$, aiming to observe celestial objects with a limiting magnitude of about 23 (5σ , point source) (H. Zhan 2021; R. Wen et al. 2024). The goal of these observations is to study the large-scale structure of the universe, dark energy, and the formation and evolution of galaxies. CSST’s main focal plane consists

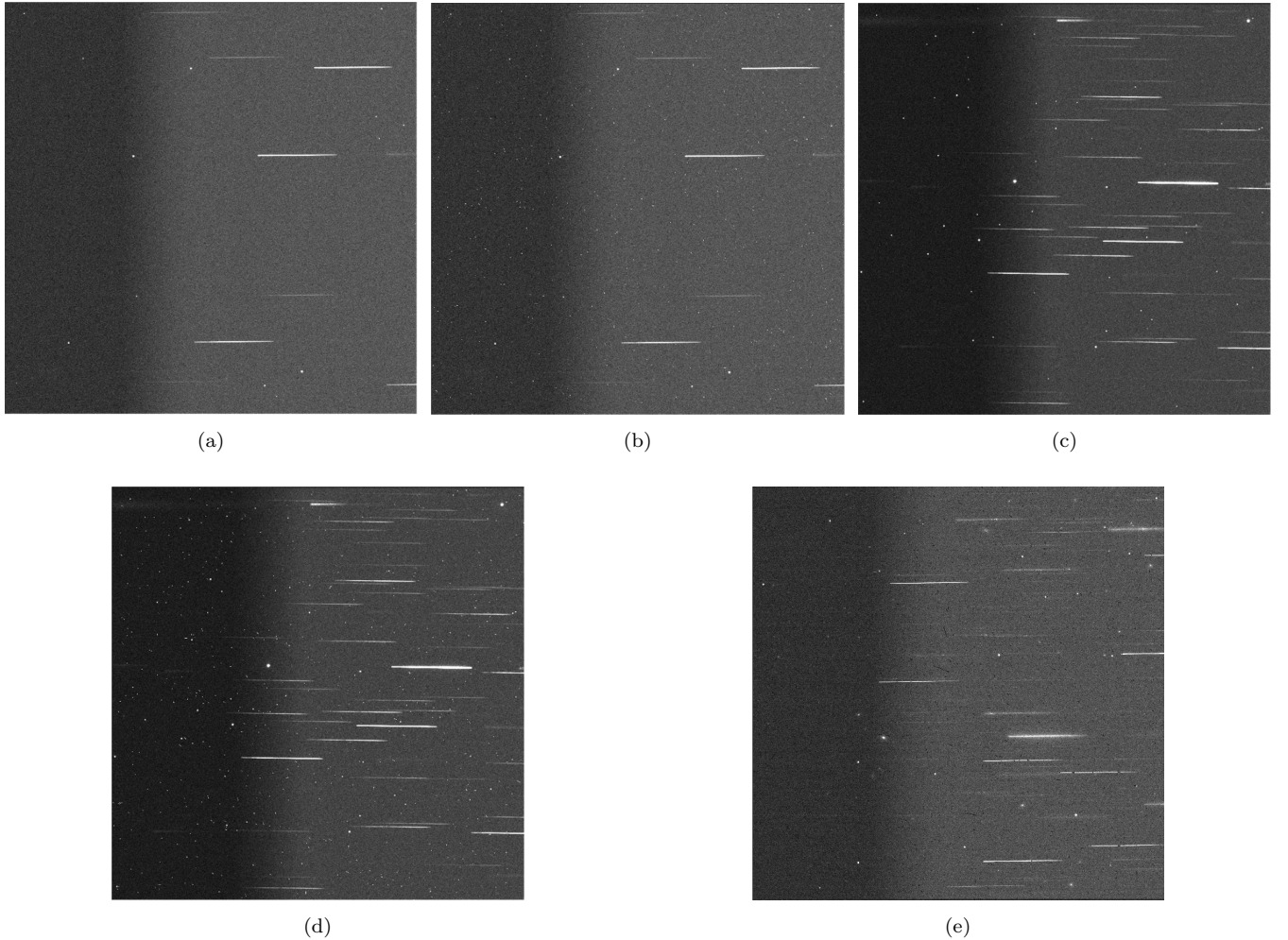


Figure 1. The simulated CSST slitless spectroscopy images, showing 1/16 of the full-size image: (a) the low-density region without cosmic rays; (b) the low-density region with cosmic rays; (c) the medium-density region without cosmic rays; (d) the medium-density region with cosmic rays; (e) the high-Galactic latitude region. The full slitless spectroscopy image has a size of approximately $9k \times 9k$ pixels. For improved detection performance, each image was divided into 16 equal sections by resampling them to 640×640 pixels. Compared to the simulated images of the Milky Way and nearby galaxy regions, the high-Galactic latitude field contains a larger number of galaxies.

of 30 detectors—18 dedicated to multiband imaging and 12 to slitless spectroscopy. Unlike HST, where detectors are switched between bands using filter and grating wheels, each detector on CSST is fixed to observe a specific band. This design eliminates the need for moving parts, reducing technical risks and improving reliability. However, due to this unique design, CSST cannot obtain photometric-slitless spectroscopy data pairs as HST does. Therefore, before processing the slitless spectroscopic data, it is essential to detect and count the sources in the slitless images, providing critical information for the subsequent spectral extraction.

Object detection is a fundamental task in computer vision, focused on identifying objects within images and determining their locations. Traditional non-deep-learning approaches rely on hand-crafted features, such as HOG (N. Dalal & B. Triggs 2005), SIFT (D. G. Lowe 2004), and Haar-like (R. Lienhart & J. Maydt 2002), and shallow models, such as the Supported Vector Machine (SVM) (C. Cortes & V. N. Vapnik 1995), AdaBoost (Y. Freund & R. E. Schapire 1997) and Deformable Part-based Model (DPM) (P. F. Felzenszwalb et al. 2010). These methods suffer from poor robustness to complex scenes, multi-scale variations, occlusions, and lighting changes, resulting in evident performance bottlenecks. In contrast, deep-learning-based approaches demonstrate clear advantages. Deep-learning methods can be broadly categorized into tow classes: Two-Stage detectors and One-Stage detectors (L. Jiao et al. 2019). Two-Stage detectors first generate region proposals and then perform classification and regression on these proposals. The advent

of deep learning, particularly the development of Convolutional Neural Networks (CNNs), has led to significant breakthroughs in this area (Z.-Q. Zhao et al. 2019). CNNs automatically learn high-level semantic features from images by combining convolutional layers, pooling layers, and fully connected layers, enabling both object classification and localization. One of the earliest frameworks to apply deep learning to object detection is R-CNN, which introduced CNNs into the field and laid the foundation for modern object-detection techniques (R. Girshick et al. 2014). Subsequent developments include Fast R-CNN (R. Girshick 2015), Faster R-CNN (S. Ren et al. 2017), and Mask R-CNN (K. He et al. 2017). Two-Stage methods achieve high accuracy and are well-suited to complex scenes; however, their multi-stage processing pipeline leads to slower inference speeds. One-Stage detectors directly predict objects from images, streamlining the computational process and thereby improving detection efficiency. They are particularly well-suited for real-time applications. YOLO (You Only Look Once) is a prominent example of One-Stage Detectors. In 2016, J. Redmon et al. (2016) introduced YOLO, treating object detection as a regression problem that directly predicts bounding boxes and class probabilities from image pixels. In addition to YOLO, notable One-Stage methods include SSD (W. Liu et al. 2016), RetinaNet (T.-Y. Lin et al. 2017), and M2Det (Q. Zhao et al. 2019). Nonetheless, these approaches generally exhibit lower accuracy on small and densely packed objects—a limitation that is particularly relevant to astronomical images, where targets are often small, closely spaced, and numerous.

Object detection in wide-field astronomical surveys presents unique challenges due to the vast areas of the sky that need to be rapidly scanned to capture a large number of celestial objects. The key to success in these surveys is the ability to quickly process massive amounts of image data. YOLO, with its ability to perform efficient and rapid object detection, is particularly well-suited for high-throughput data environments, meeting the real-time performance and resource utilization demands of such survey projects. As a result, the application of YOLO in astronomical object detection has been increasing. Researchers have successfully applied YOLO models to detect galaxies (Y. Xing et al. 2023; F. Wang et al. 2023), galaxy clusters (K. Grishin et al. 2025), solar radio bursts (SRBs) (E. Carley 2020), and streaks caused by moving objects like satellites or meteorites (L. Varela et al. 2019). Additionally, YOLO has shown strong performance across various observational datasets, including photometric, slitless spectroscopic and radio data (Z. He et al. 2021; D. Cornu et al. 2024; Y. Zhou et al. 2025; A. Anthore & D. Cornu 2024). Traditionally, object detection in astronomy often relied on supplementary data, such as photometric catalogs or photometric redshift information, to detect galaxy clusters. In contrast, YOLO’s ability to work directly on images allows for faster, more convenient detection while avoiding the systematic errors often associated with traditional methods (K. Grishin et al. 2025).

This study focuses on identifying zero-order images and first-order spectral lines of stars and galaxies in slitless spectroscopic data using the YOLO algorithm, addressing the unique challenges posed by high target densities in astronomical images. The structure of this paper is organized as follows. Section 2 provides an overview of the training and test samples used to construct our object detection model. The methods employed in this study are detailed in Section 3. Section 4 presents the detection results and a detailed discussion is presented in Section 5. Finally, we summarize and conclude our study in Section 6.

2. DATA

2.1. Slitless Spectroscopic Data of CSST

The CSST (H. Zhan 2021) is a large space-based optical telescope with a 2-meter aperture, equipped with a variety of advanced observational instruments, including a wide-field optical survey module, a terahertz module, a multi-channel imager, an integral field spectrograph, and an exoplanet imaging coronagraph. Among these instruments, the wide-field optical survey module is one of the most critical observational devices of the CSST. The main focal plane of the wide-field optical survey module covers an area of approximately $1.1^\circ \times 1.2^\circ$ at the center of the field of view and consists of 30 detectors. Among these 30 detectors, 18 are dedicated to multi-band imaging observations (NUV, u, g, r, i, z, y), while 12 are assigned to slitless spectroscopic observations (GU, GV, GI). Each exposure generates 30 observational images, corresponding to both the multi-band imaging and the slitless spectroscopic images. Each detector used for slitless spectroscopic observation by the CSST corresponds to two gratings, meaning each grating unit in the GU, GV, and GI bands consists of two gratings. The spectra from these two gratings, which operate at different orders (+1 or -1), disperse in opposite directions. In the final slitless spectroscopic image, the demarcation line is located at the $2/5$ position of the detector, with sources on the left dispersing towards the right and sources on the right dispersing towards the left. In slitless spectroscopic images, our primary detection targets are the zero-order images and first-order spectra of stars and galaxies. The zero-order spectrum is undispersed and exhibits better image

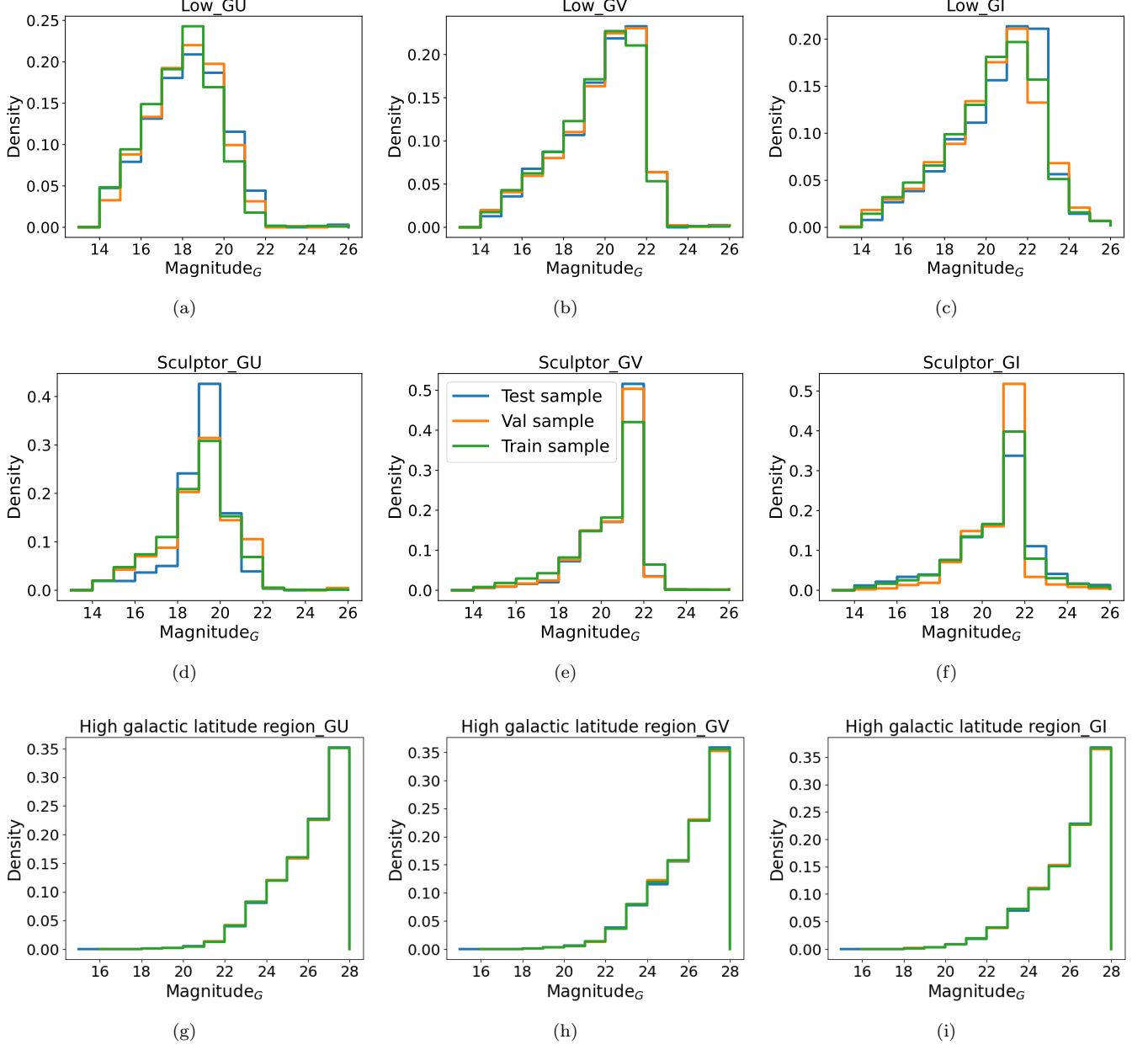


Figure 2. The normalized distribution of datasets in the GU, GV, and GI bands for high galactic latitude regions, as well as for low-density and medium-density regions of the Galactic and nearby galaxies, where the vertical axis “Density” represents the proportion of sources within that magnitude interval relative to the total number. For the Milky Way and nearby galaxy regions, the simulated images use g-band magnitudes, whereas for the high-Galactic latitude region, the slitless spectroscopy images use white-light magnitudes corresponding to each chip. The distributions for the training, validation, and test sets within each chip and region are generally consistent, ensuring balanced training performance.

quality, accounting for approximately 10% of the total incident energy. It serves as a reference point for wavelength calibration. The first-order spectrum, which is the primary dispersed spectral signal, contains wavelength-dependent information.

2.2. Data Simulation

To ensure the timeliness and reliability of scientific outputs from the CSST, researchers have developed a comprehensive data simulation software designed to evaluate data processing pipelines and scientific analysis tools. Within this

framework, the slitless spectral simulation module is built upon a detailed model of the optical system, incorporating the diffraction characteristics of the grating to achieve realistic and accurate results. The simulated slitless spectroscopic data incorporate realistic observational characteristics by modeling the relationship between spectral traces and wavelength through polynomial fitting, while accounting for instrumental effects such as the point spread function (PSF), detector noise, and charge transfer efficiency. The simulations are generated using astrophysical source catalogs combined with optical response functions and detector characteristics to produce realistic slitless spectral images (X. Zhang et al. 2025; C.-L. Wei et al. 2025).

In this study, we used Cycle 6 (C. Wei et al. 2023), the software for simulating CSST Main Sky Survey data, to generate synthetic CSST slitless spectroscopy images of the Galactic and nearby galaxies. The simulated data products for Cycle 6 are centered at $Ra = 244.9727$ degrees, $Dec = 39.8959$ degrees, covering approximately 1.53 square degrees. They include 137 continuously overlapping exposure pointings. There is a shift of approximately one CCD size between consecutive exposures, covering a total sky area of about 2.65 square degrees. The central 1.53 square degree region constitutes the complete survey area. These simulated data primarily represent stars, with galaxies excluded for this simulation. For practical applications, it is assumed that the data input into the model has already undergone correction for instrumental effects. However, it cannot be guaranteed that cosmic ray contamination has been fully removed. To account for this, cosmic ray effects were added to half of the simulation data, while the other half only included sky background noise. For the data with cosmic ray effects, different cosmic ray maps were generated for various outputs (e.g., calibrated output CAL and main science output MS) during the simulation process. This approach helps to realistically simulate the impact of cosmic rays on the detector, as their distribution is random, and different observational stages may be affected by different cosmic ray events. Two input star catalogs were selected: one for a low-density celestial region and the other for a medium-density region, specifically the Sculptor region. These two input catalogs are stellar catalogs, consisting of stars selected from the same sky region in Gaia DR3. They retain the spatial coordinates provided in Gaia DR3, along with the three Gaia magnitudes: G, BP, and RP. A star-by-star correspondence is established by comparing the G, BP, and RP magnitudes of stars in Gaia DR3 with those generated by the Galaxia software. In Galaxia, stellar magnitudes are computed via synthetic photometry: the Spectral Energy Distributions (SEDs), produced using the BT-Settl library based on stellar parameters (T_{eff} , $\log g$, $[Fe/H]$), are convolved with the CSST filter transmission curves to derive magnitudes in each band (C.-L. Wei et al. 2025). For each Gaia DR3 star S_GDR3, a corresponding Galaxia star S_Gal with consistent G, BP, and RP magnitudes is then selected. The parameters and ugriz magnitudes of S_Gal are then assigned to S_GDR3 to facilitate matching with model spectra. Since all stars in the Gaia DR3 sample are brighter than $G=21$ mag, the stars in Gaia DR3 are used to replace those with $G<21$ mag in the simulated Galaxia catalog. These star catalogs are used for the development of the pipeline in the process of positional and flux calibration, as well as for other CSST data processing tasks (Y. Zhou et al. 2025). These catalogs contain detailed information on stars within their respective regions. For stars with a Gaia G magnitude less than 22, the density in the low-density region is 5,197.76 objects per square degree, while in the medium-density region, it is 7,328.82 objects per square degree. The simulation program generates simulated images based on a pointing file (pointing.data). The survey module's focal plane consists of 30 detectors, and each pointing generates 30 images per exposure (150 seconds). Among these, 12 are slitless spectroscopy images. The dataset includes 30 exposure pointings for both the low-density and medium-density celestial regions, with each pointing consisting of two types of images: those with cosmic rays and those without. Additionally, each pointing includes 12 slitless spectroscopy images in the GU, GV, and GI bands. As a result, this sample set contains a total of 1,440 slitless spectroscopy images, with 720 images from the low-density celestial region and 720 from the medium-density celestial region. Figure 1 (a)-(d) presents four simulated CSST slitless spectroscopy images of the Galactic and nearby galaxies. To intuitively demonstrate the contamination of cosmic rays on astronomical images and their distinction from real astrophysical signals, Figure 1 provides a comparative set of simulated image pairs. Panels (a) and (b) show slitless spectroscopic images of the same low-density region in the Galactic and nearby galaxies. Specifically, (a) is free of cosmic rays, while (b) includes cosmic rays—all other content remains identical. Similarly, panels (c) and (d) present slitless spectroscopic images of the same medium-density region in the Galactic and nearby galaxies, where (c) is without cosmic rays and (d) includes them—again, all other elements are unchanged. Cosmic rays typically appear as sharp, bright peaks confined to one or a very small cluster of pixels, exhibiting a morphology inconsistent with the telescope's point spread function (PSF). In contrast, genuine astrophysical point sources display smooth, symmetric profiles that conform to the expected PSF shape.

Table 1. The observation numbers for the GU, GV, and GI bands.

Region	GU bands	GV bands	GI bands
the low-density regions	35,266	34,440	34,122
the medium-density regions	54,416	50,588	51,516
the high galactic latitude regions	25,369	27,756	46,033

Note: The table displays the count of sources with magnitudes brighter than 22 in the output star catalog. The output catalog is a source information file produced alongside the simulated images by the simulation software Cycle 6 and Cycle 9.

For the high galactic latitude cases, we used the Cycle 9 simulation data (C. Wei et al. 2024), which includes a larger number of galaxies. The simulation data products for Cycle 9 comprise five simulated sky regions, centered at $[Ra, Dec] = [180.0, 60.0], [240.0, 30.0], [170.0, -23.0], [30.0, -48.0],$ and $[300.0, -60.0]$, respectively. Each region includes simulated images and slitless spectral data for galaxies and stars, covering an area of 5 square degrees, with a total of 1,221 exposure pointings. Furthermore, to account for the effect of intrinsic ellipticity in current weak gravitational lensing shear measurements, the intrinsic ellipticity of each galaxy within the regions was rotated by 90 degrees in the Cycle 9 simulations before the imaging process was repeated. In total, image simulations for 50 square degrees were completed. In each continuous sky region, there is a shift of roughly one CCD size between consecutive exposures. The central 5 square degree area represents the complete survey region. We selected 10 exposure pointings from the Cycle 9 simulation data, and Figure 1 (e) shows a simulated CSST slitless spectroscopy image of the high galactic latitude region.

Finally, the complete dataset consists of 1,560 slitless spectroscopy images: 1,440 from Cycle 6 contain only stars and no galaxies, while the remaining 120 from Cycle 9 include both stars and a substantial population of galaxies. Figure 2 illustrates the distribution of datasets across the GU, GI, and GV bands for the low-density and medium-density regions of the Galaxy and nearby galaxies, as well as for the high galactic latitude regions. The distributions for the training, validation, and test sets are generally consistent. And the observation numbers for the GU, GV, and GI bands are shown in Table 1.

Our dataset contains only ordinary stars and galaxies, and does not contain special types such as AGNs. Although AGNs are not explicitly included, this does not affect our results, as the aim of this work is to estimate the number and counts of sources. In slitless spectroscopy, both AGNs and typical galaxies appear as extended-source spectra, and thus their distinction is not critical for our model’s learning process.

3. METHODOLOGY

3.1. YOLO

In this work, we developed a model to detect zeroth-order images and first-order spectra in simulated CSST slitless spectroscopy images using the YOLOv8 algorithm, a member of the YOLO family (J. Redmon et al. 2016). The YOLO algorithm is a groundbreaking deep learning approach for real-time object detection. It reformulates the object detection task as a regression problem, directly predicting bounding box coordinates and class probabilities from image pixels, which enables fast and efficient detection. YOLO divides the input image into an $S \times S$ grid, where each grid cell predicts B bounding boxes along with their confidence scores. Each grid cell also predicts C conditional class probabilities. The final output is a tensor of size $S \times S \times (B \times 5 + C)$, where $B \times 5$ corresponds to the bounding box coordinates (x, y , width w , height h) and confidence scores, while C represents the class probabilities. The YOLO loss function consists of three components: bounding box coordinate loss, confidence loss, and class probability loss (J. Redmon et al. 2016). These losses are combined in a weighted sum for multi-objective optimization. YOLO processes images in real-time, and because it observes the entire image during both training and detection, it can use global contextual information to minimize false positives from background noise. However, YOLO may encounter difficulties in localizing small or densely packed objects.

YOLOv8, released by Ultralytics on January 10, 2023, introduces significant advancements over its predecessors. It features a deeper network architecture with additional convolutional layers and residual connections, which enhance the network’s capacity for feature extraction and improve detection performance. YOLOv8 also offers models of varying

sizes, including YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l, and YOLOv8x, allowing for flexible applications based on computational requirements. For this project, we selected the large-scale architecture, YOLOv8l, from the Ultralytics YOLOv8 library to develop an object detection model aimed at identifying zeroth-order images and first-order spectra in simulated CSST slitless spectroscopy images.

3.2. Data Preprocessing

Due to the high density of sources in some of the slitless spectroscopy images for the medium-density celestial region (as shown in Figure 1 (c) and (d)), detecting objects using deep learning poses significant challenges. To address this, we divided each slitless spectroscopy image into 16 equal parts by resampling each section to 640×640 pixels, as the YOLOv8 model requires input images of this size. This approach resulted in a total of 24,960 segmented slitless spectroscopy images, each of 640×640 pixels, for training and testing. These images were annotated using CVAT (B. Sekachev et al. 2019), a interactive video and image annotation tool for computer vision, to accurately define the bounding box dimensions, which enclose the detected objects. Each label includes the object's central coordinates (x , y) and the width and height of the bounding box (w , h). The labels were visually annotated, and the completeness of visual annotation is shown in Figure 3. The final dataset of 24,960 segmented images was split into training, validation, and test sets at a ratio of 8:1:1. Specifically, 19,968 images were used for training, 2,496 images for validation, and 2,496 images for testing.

3.3. Model Training

During model training, a batch size of 16 was used, and the Stochastic Gradient Descent (SGD) optimizer was employed to fine-tune the model parameters. The parameters for image translation, image scaling, and image mosaicking were set to 0.1, 0.5, and 1.0, respectively. The initial learning rate was configured to 0.01, with the Intersection over Union (IoU) training threshold set to 0.7. After 920 epochs of training, completed in 206.911 hours, the model's loss showed stable convergence. As a result, we successfully developed an object detection model capable of identifying zeroth-order images and first-order spectra in the simulated CSST slitless spectroscopy images.

To evaluate the model's performance in detecting sources, we computed the precision and recall for the test dataset images. These metrics were then used to calculate the model's mAP score. The formulas for computing precision, recall, average precision (AP), and IoU are as follows:

$$precision = \frac{TP}{TP + FP} \quad (1)$$

$$recall = \frac{TP}{TP + FN} \quad (2)$$

$$AP = \int_0^1 P(R) dR \quad (3)$$

$$IOU = \frac{A \cap B}{A \cup B} \quad (4)$$

where TP (True Positives) refers to the number of targets correctly identified, FP (False Positives) denotes the number of targets incorrectly identified, and FN (False Negatives) represents the number of targets that were missed or incorrectly rejected. A represents the predicted bounding box, while B represents the ground-truth bounding box. Precision and recall are key metrics for evaluating the performance of machine learning models. Precision measures the ratio of true positives (correctly detected objects) to the total number of detected objects, reflecting the accuracy of the detection results, while recall measures the ratio of true positives to the total number of actual existing objects, reflecting the completeness of the detection method (L. Jiao et al. 2019). Typically, precision and recall exhibit a trade-off: improving one often leads to a reduction in the other. To balance these metrics, the AP is used, which represents the area under the Precision-Recall curve. The AP value ranges from 0 to 1, with a higher value indicating better model performance. The mAP is a metric used to evaluate the performance of object detection models by calculating the average of the AP scores across multiple object categories. mAP@50 (mAP50) is a composite metric that averages the AP scores across multiple object categories, providing an overall performance evaluation of the detection model. mAP@50 (also known as mAP50) considers only those detections where the IoU is at least 50%, while mAP50-95 calculates the mAP considering IoU thresholds ranging from 50% to 95%.

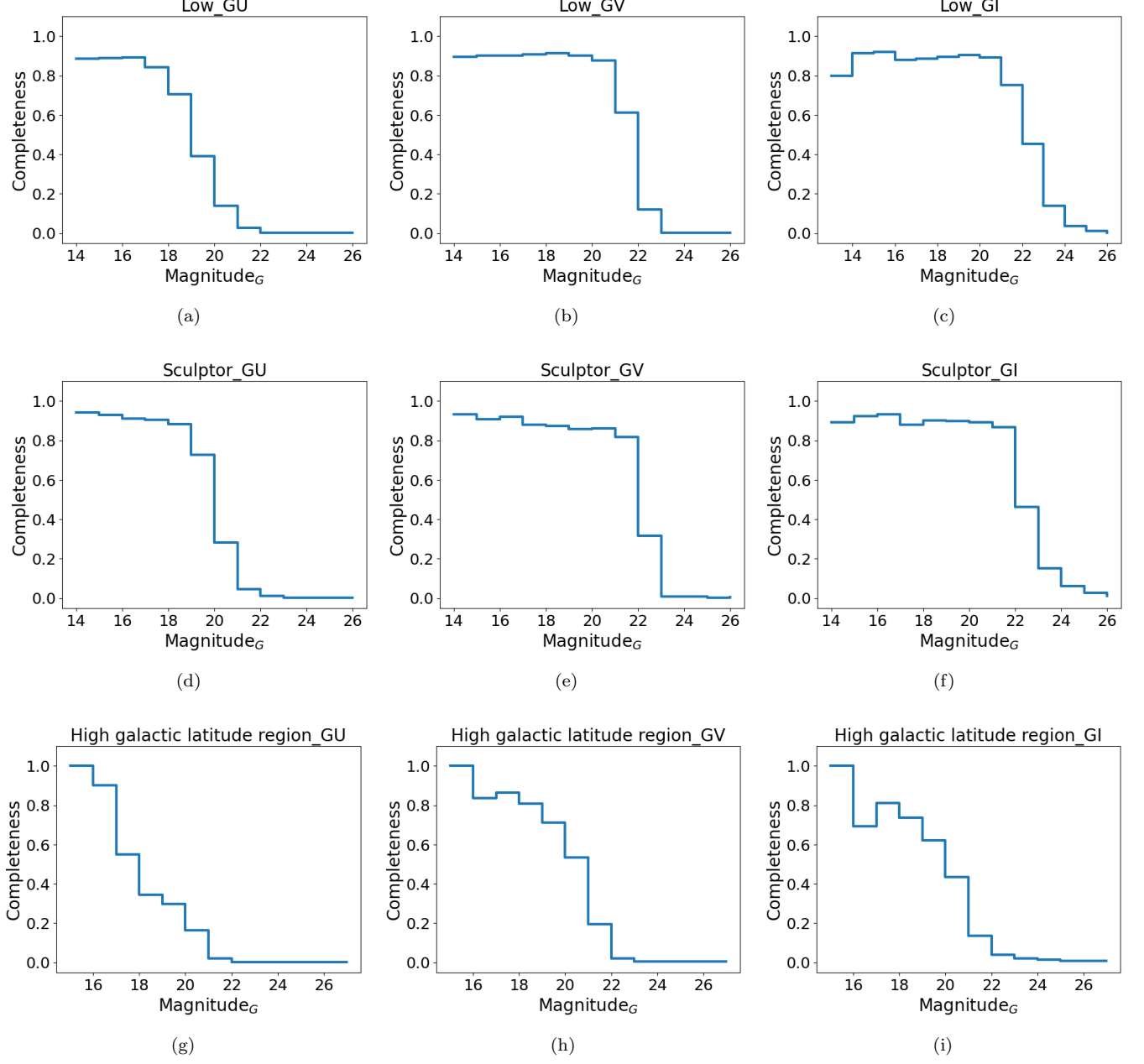


Figure 3. The completeness of visual annotation. The completeness of manual annotation is lowest in the high galactic latitude regions, primarily because these regions contain a large number of galaxies, which are extended sources. At a given magnitude, galaxies appear visually fainter than point sources, reducing their detectability. Moreover, the completeness varies across different bands due to differences in SNRs. In particular, the GU band images exhibit relatively low SNR, which limited the number of sources that could be reliably identified through visual inspection.

4. RESULTS

The detection model outputs the center coordinates of the object, along with the width and height of the bounding box, and the confidence score. The confidence level represents the probability that an object is present within a given bounding box. In our study, we used a standard confidence threshold of 0.001 during validation. This resulted in precision values of 87.8%, 88.6%, and 87.0%, and recall rates of 85.6%, 90.4%, and 80.8% for the total, spectral lines, and zeroth-order images, respectively. The detailed test results are presented in Table 2. Figure 4 shows the precision-recall curve for the detection model. The performance for detecting spectral lines is superior to that for

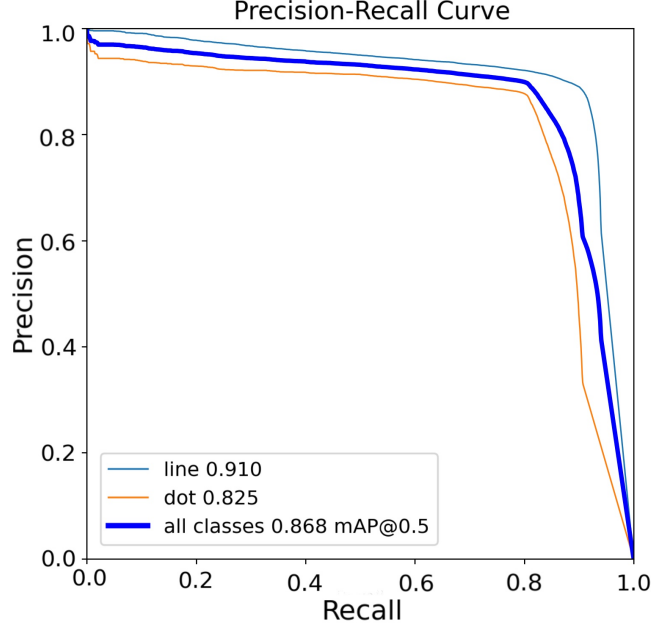


Figure 4. The precision-recall curve of the source detection model, evaluated on the validation dataset. The model achieves a mean Average Precision (mAP@0.5) of 0.868 across all classes, with the “line” class (spectral lines) performing best at 0.910 precision and the “dot” class (zero-order image) at 0.825.

Table 2. Performance Metrics of Slitless Spectroscopy Source Detection Model Across Different Classes

Class	Precision	Recall	mAP50	mAP50-95
all	0.878	0.856	0.868	0.435
spectral lines	0.886	0.904	0.910	0.515
zeroth-order images	0.870	0.808	0.825	0.354

zeroth-order images. This may be attributed to the influence of cosmic rays, which can be confused with zeroth-order images. Additionally, the SNR for zeroth-order images is lower, as their energy contributes only about 10% of the total incident energy (H. Zhan 2021).

In the test set, we used a standard confidence threshold of 0.07. Each image was processed in approximately 9.3 ms, resulting in a total processing time of 0.387 minutes for all 2,496 images.

Figure 5 presents the schematic diagrams of slitless spectroscopy image detection. To accurately assess the detection rates for different magnitudes in the test set, we matched the coordinates of the zeroth-order images with the source coordinates from the star catalog. The star catalog provides the coordinates from direct imaging. Using the following dispersion function (X. Zhang et al. 2025; C.-L. Wei et al. 2025), the position coordinates of the zeroth-order images and spectral lines were determined:

$$\begin{aligned} \Delta y(x, y, \Delta x) = & \\ & (b_{00} + b_{01}x + b_{02}y + b_{03}x^2 + b_{04}xy + b_{05}y^2) + \\ & \Delta x(b_{10} + b_{11}x + b_{12}y + b_{13}x^2 + b_{14}xy + b_{15}y^2) \end{aligned} \quad (5)$$

the position coordinates of the zeroth-order images and spectral lines can be obtained. Here, x and y represent the coordinates from direct imaging, while $x + \Delta x$ and $y + \Delta y$ correspond to the coordinates of the zeroth-order image and first-order spectrum in the detected image. Figure 6 shows the detection rates of sources at different magnitudes

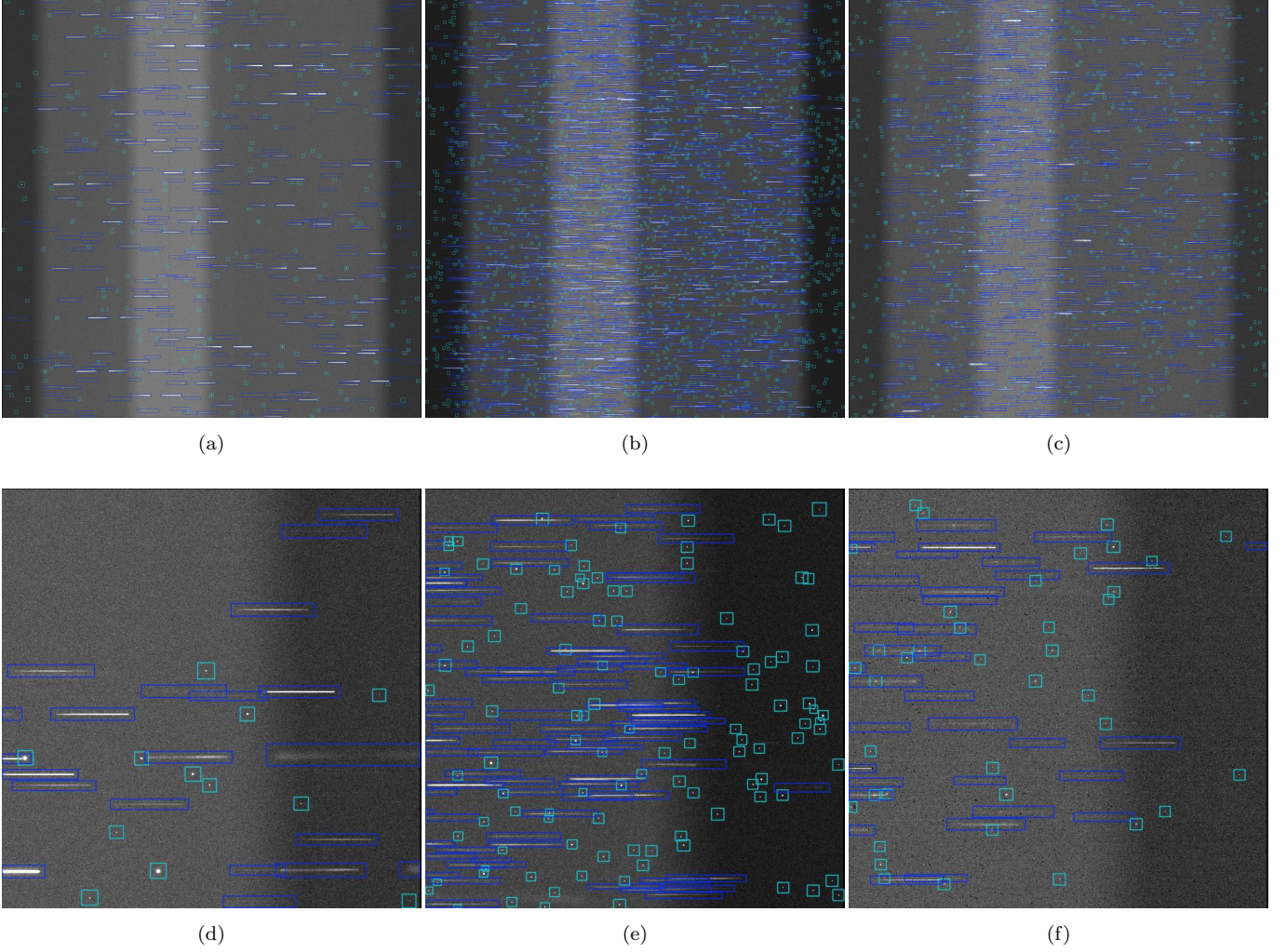


Figure 5. Schematic diagram of slitless spectroscopy image detection: (a) Illustration of target detection in a low-density celestial region of the Galactic and nearby galaxies, (b) Illustration of target detection in a medium-density celestial region of the Galactic and nearby galaxies, (c) Illustration of target detection in a high galactic latitude region. The figure shows the final CSST slitless spectral simulation image, assembled by stitching together 16 sub-images after object detection. Panel (d), (e), and (f) are local images magnified 16 times from panel (a), (b), and (c) respectively, to better demonstrate the target detection effect.

in the GU, GI, and GV bands. More details can be found in Tables 3, 4, and 5. Excluding the low-SNR images in the GU band, the model achieved detection rates exceeding 80% for targets brighter than 21 mag in the GI and GV bands within medium-density regions of the Galactic and nearby galaxies. In low-density regions of these fields, detection rates remained above 80% for targets brighter than 20 mag in the GI and GV bands. For the high galactic latitude region, the model achieved detection rates exceeding 70% for targets brighter than 18 mag in the GI and GV bands. And the median location distance of the zero-order images between the catalog and the YOLO output is shown in Figure 7. The detection rate and the median location distance presented here were obtained by matching the coordinates of the zero-order images output by the model with the coordinates of the zero-order images provided in the catalog (specifically, by converting the coordinates of the direct imaging given in the catalog using Equation (5)). We did not evaluate the detection rate or median location distance for spectral lines because their positions and bounding boxes, derived from the star catalog and Equation (5), are affected by field distortions. These distortions vary across the image, lack a precise analytical description, and are instead approximated through interpolation of optical simulation data, with maximum pixel offsets reaching up to 200 pixels (X. Zhang et al. 2025; C.-L. Wei et al.

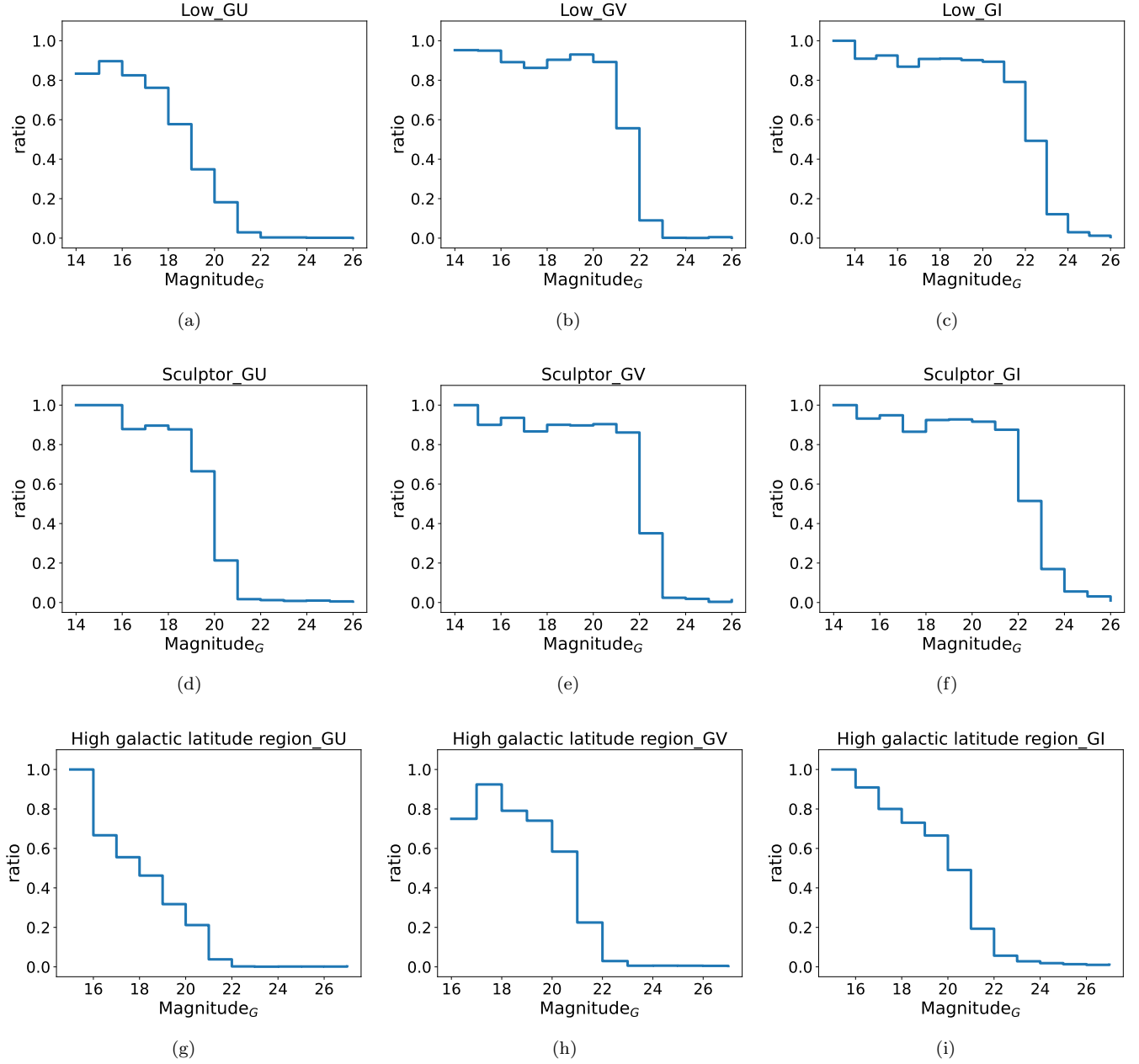


Figure 6. The detection rates of different magnitudes in the GU, GI, and GV bands for high galactic latitude regions, as well as low-density and medium-density regions of the Galactic and nearby galaxies.

2025). Therefore, we cannot obtain accurate magnitude values and positional coordinates for the spectral lines as we can for the zero-order images.

In the CSST data processing pipeline, the preprocessing stage incorporates cosmic ray removal procedures. However, complete elimination cannot be guaranteed. Consequently, our training set intentionally includes some images with cosmic rays. These cosmic rays exhibit morphological similarities to zero-order images in slitless spectral data (though with different PSF profiles), making them potentially susceptible to confusion with actual zero-order signals. Furthermore, we have conducted comparative evaluations on the test set to analyze the impact of cosmic rays on model performance specifically regarding zero-order image detection (see Table 6). We can observe that the presence of cosmic rays slightly degrades the model’s performance.

Table 3. The number of cataloged and detected sources in low-density region.

Mag	GU			GI			GV		
	Cat	Pre	Rates	Cat	Pre	Rates	Cat	Pre	Rates
13	0	0	–	2	2	100%	0	0	–
14	72	60	83.33%	44	40	90.91%	42	40	95.24%
15	116	104	89.66%	134	124	92.54%	120	114	95.00%
16	200	165	82.5%	198	172	86.87%	240	214	89.17%
17	298	227	76.17%	304	276	90.79%	320	276	86.25%
18	452	261	57.74%	474	431	90.93%	374	338	90.37%
19	608	212	34.87%	572	516	90.21%	574	534	93.03%
20	726	132	18.18%	810	724	89.38%	762	680	89.24%
21	1300	38	2.92%	1218	964	79.15%	1162	647	55.68%
22	1716	6	0.35%	1868	921	49.30%	1802	162	8.99%
23	0	0	–	1856	225	12.12%	1768	3	0.17%
24	2116	4	0.19%	1982	59	2.98%	2084	2	0.1%
25	0	0	–	2440	30	1.23%	2380	12	0.5%
26	3074	1	0.03%	3140	17	0.54%	2916	5	0.17%

Table 4. The number of cataloged and detected sources in medium-density region.

Mag	GU			GI			GV		
	Cat	Pre	Rates	Cat	Pre	Rates	Cat	Pre	Rates
14	70	71	100%	48	48	100%	52	52	100%
15	72	72	100%	88	82	93.18%	80	72	90.0%
16	148	130	87.84%	136	129	94.85%	140	131	93.57%
17	212	190	89.62%	178	154	86.52%	180	156	86.67%
18	968	849	87.71%	318	294	92.45%	622	560	90.03%
19	2070	1376	66.47%	552	512	92.75%	1284	1152	89.72%
20	2268	483	21.3%	680	623	91.62%	1474	1332	90.37%
21	7690	130	1.69%	1484	1299	87.53%	4564	3931	86.13%
22	832	10	1.2%	832	428	51.44%	764	268	35.08%
23	998	8	0.8%	956	162	16.95%	918	22	2.4%
24	1128	11	0.98%	1254	70	5.58%	1086	20	1.84%
25	1628	9	0.55%	1590	49	3.08%	1496	5	0.33%
26	1918	9	0.47%	1954	20	1.02%	1760	21	1.19%

5. DISCUSSION

In this section, we discussed the performance of our source detection model for CSST slitless spectroscopy. We discussed dataset-related considerations, evaluated the impact of image segmentation strategies, identified key challenges such as low-SNR regions and extended galaxy morphology that lead to missed detections, and assessed the model’s robustness against a variable ratio of zeroth-order images to spectra. We also compared the performance differences between our method and traditional approaches.

Table 5. The number of cataloged and detected sources in the high Galactic latitude region.

Mag	GU			GI			GV		
	Cat	Pre	Rates	Cat	Pre	Rates	Cat	Pre	Rates
15	1	1	100%	1	1	100%	0	0	–
16	3	2	66.67%	11	10	90.91%	4	3	75.0%
17	45	25	55.56%	65	52	80.0%	66	61	92.42%
18	80	37	46.25%	204	149	73.04%	153	121	79.08%
19	214	68	31.78%	508	338	66.54%	320	237	74.06%
20	515	109	21.17%	1125	552	49.07%	685	400	58.39%
21	1629	62	3.81%	2396	463	19.32%	1629	366	22.47%
22	4073	9	0.22%	4727	268	5.67%	4153	122	2.94%
23	6955	7	0.1%	7820	220	2.81%	7068	41	0.58%
24	9448	15	0.16%	11371	213	1.87%	9669	59	0.61%
25	12899	21	0.16%	16327	222	1.36%	13641	79	0.58%
26	19132	29	0.15%	25416	259	1.02%	20375	103	0.51%
27	13609	25	0.18%	18165	193	1.06%	14636	63	0.43%

Table 6. The Impact of Cosmic Ray on Model Performance Across Different Bands

Band	With CR			Without CR		
	Precision	Recall	F1-Score	Precision	Recall	F1-Score
GU	0.73	0.23	0.35	0.95	0.29	0.44
GV	0.85	0.82	0.84	0.98	0.86	0.92
GI	0.84	0.87	0.85	0.96	0.88	0.92

Note:

1. *With CR* represents images containing cosmic rays, and *Without CR* represents images without cosmic rays.
2. The recall in the GU band is very low because the GU band images have a very low SNR, making it difficult for the model to identify sources.

Our dataset comprises three distinct sky regions—high galactic latitude fields (containing both galaxies and stars), medium-density fields at the Galactic and nearby galaxies (containing only stars), and low-density fields at the Galactic and nearby galaxies (containing only stars). The simulated data for low-density fields at the Galactic and nearby galaxies represent the most fundamental and easily identifiable scenario. Beyond this, we aimed to enable the model to handle regions with significant spectral crowding, which led us to incorporate simulated data from medium-density fields at the Galactic and nearby galaxies. Finally, to ensure the model can also identify galaxies, we included simulated data from high galactic latitude fields, as these regions contain a substantial population of galaxies. Beyond stars and galaxies, rarer objects such as nebulae and active galactic nuclei may also be present—objects not considered in the current study. Although our model does not include AGNs, since our work focuses on the estimation of source numbers and counts, whether it is an AGN or a typical galaxy, it actually manifests as an extended source spectrum in slitless spectroscopy. This does not affect our learning outcomes, but in real observations, such a model still needs to be tested. Future work will need to further refine the dataset by incorporating galaxies across different sky regions, additional rare object types, and a wider range of density conditions.

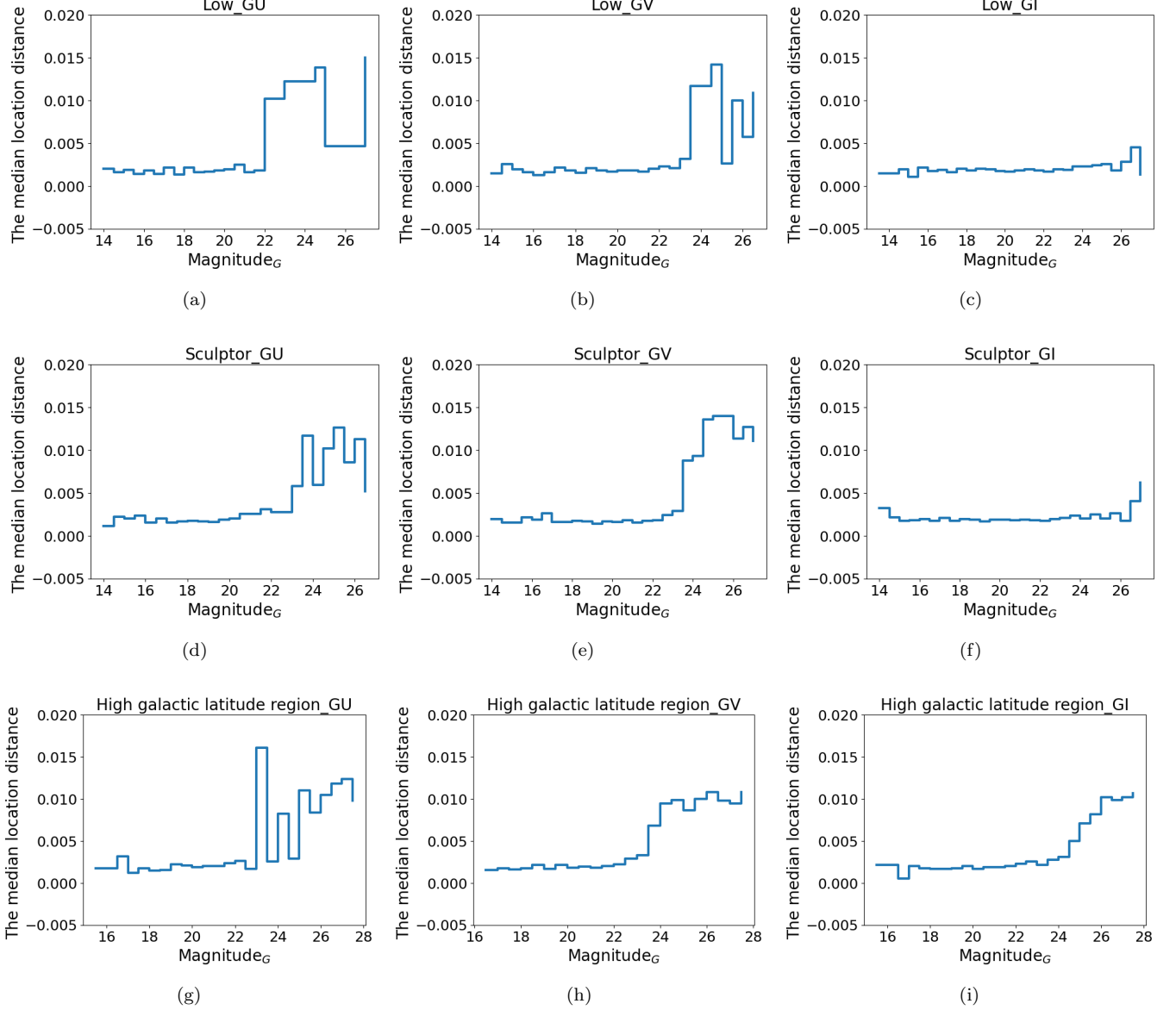


Figure 7. The median location distance of the zero-order images between the catalog and the YOLO output. The median location distance is measured in normalized relative coordinates. Given an image size of 640×640 pixels, a distance of 0.01 indicates that the positions of the zero-order images differ by 6.4 pixels between the catalog and the model output.

During training and testing, each slitless spectroscopy image was divided into 16 equal sections by resampling them to 640×640 pixels. This approach can cause some spectra to be truncated at the segment boundaries, potentially leading to partial recognition (e.g., only half of a spectrum) when the segmented images are reassembled into a complete image along the junction lines. To address this issue, some researchers have adopted a method of horizontal slicing followed by vertical stretching to avoid such boundary artifacts (Y. Zhou et al. 2025). The authors first slice the original image into ten horizontal strips with a 62-pixel overlap between neighboring strips, preventing targets from being truncated exactly at the edges. Each strip is then stretched vertically to a square format to meet YOLO’s input requirement. Since the spectra are elongated horizontally, this horizontal slicing preserves their full extent and thereby eliminates boundary artifacts. However, stretching cropped images into square dimensions can introduce geometric distortions, potentially leading to errors in subsequent spectral extraction or morphological analysis. In the case of zero-order images, cropping and resizing may be particularly unsuitable, as their inherently small size increases the risk of losing

critical details. Future work should therefore consider adapting cropping strategies to the specific scientific objectives of each analysis.

In slitless spectroscopy images, in addition to the first-order spectra, higher-order spectra such as second-order spectra occasionally appear. However, the first-order spectra generally have much higher energy than the higher-order ones. Our goal is to detect the zero-order images and first-order spectra of stars and galaxies. Typically, a single “source” manifests as one zero-order image and one first-order spectrum in the slitless spectroscopic images. We need to count the number of sources (i.e., the number of zero-order images) and determine the positions of both the zero-order images and their corresponding first-order spectra generated by these sources. When using CVAT for annotation, we primarily label zero-order images and first-order spectra, without annotating spectral lines of other orders. For most sources, the majority of energy is concentrated in the first-order spectra, while spectral lines of other orders are typically not visible. Although brighter sources may exhibit visible spectral lines of other orders, these are usually distinct from first-order spectra—spectral lines of other orders tend to be broader and easier to distinguish. Nevertheless, occasional confusion between first-order and higher-order spectra may have occurred, leading to annotations that include not only first-order spectra but also a small number of second-order spectra. A similar issue may arise during detection as well. Therefore, when calculating the number of sources and detection rates, we used the detected zeroth-order images as the basis for the statistics.

The energy of zeroth-order images accounts for only about 10% of the total incident energy (H. Zhan 2021), which results in a lower SNR for zeroth-order images compared to spectral lines. The lower SNR can be attributed to the fact that the signal strength of zeroth-order images is weaker, making them more susceptible to noise interference. Zeroth-order images, though characterized by lower energy and SNR, are indispensable for wavelength calibration as they provide reliable reference positions. Wavelength calibration is a vital process in spectroscopy, as it ensures the accurate determination of the wavelengths of detected spectral lines. Without zeroth-order images, the detected spectral lines would lack the necessary reference for precise wavelength calibration, rendering them essentially useless for further analysis and interpretation.

We need to count the number of sources (i.e., the statistical count of zero-order images), but the low energy of zeroth-order images may render some of them undetectable in slitless spectral images, resulting in variability in the ratio of visible zeroth-order images to spectral lines. To evaluate the model’s detection performance of zero-order images under different ratios of zeroth-order to spectral components, we selected a GV-band slitless spectral image from a medium-density Galactic and nearby galaxy region for testing. The image was divided into 16 equal segments. For each sub-image, we randomly selected multiple sources and masked their corresponding zeroth-order images by replacing the affected pixels with the median values of surrounding pixels (Figure 8), followed by source detection analysis. The surrounding pixels are defined by dilating the original target bounding box—which is derived by applying Equation (5) to convert the direct imaging coordinates from the catalog file—with a 1-pixel kernel, thus creating an expanded region that encompasses both the original bounding box surrounding the point target and its immediate one-pixel border. Figure 11 shows the detection rates of zero-order images brighter than the 22 magnitude at various masked ratios. The results indicate that lowering the ratio of zeroth-order images to spectral lines has little effect on the detection performance of zero-order images, demonstrating the model’s robustness and stability under varying conditions. However, in practical applications, it is still essential to strike a balance between the number of zeroth-order images and spectral lines to ensure reliable wavelength calibration.

Each detector used by the CSST for slitless spectroscopic observations is paired with two gratings that disperse light in opposite directions. As a result, the SNR of targets tends to be lower at the grating junctions in the final slitless spectroscopic image. Visual inspection of the detection results reveals that, most missed detections occur at these junctions, especially for slitless spectroscopic images of medium-density regions within the Galactic and nearby galaxies, as illustrated in Figure 9. And the missed detections account for approximately 12.8% of the total visual inspection.

In slitless spectroscopic images of high galactic latitude regions, galaxies constitute a large proportion of the observed sources. In the test set, the ratio of the number of stars to galaxies with magnitude values less than 22 is approximately 0.072. Unlike stars, galaxies exhibit more diffuse light distributions across the field of view. According to the SNR formula (L. Mortara & A. Fowler 1981; C. A. Gullixson 1992):

$$\frac{S}{N} = \frac{S_{\star}}{\sqrt{S_{\star} + n_p \cdot \left(1 + \frac{n_p}{n_s}\right) \cdot (S_S + t \cdot dc + R^2 + G^2 \sigma_f^2)}} \quad (6)$$

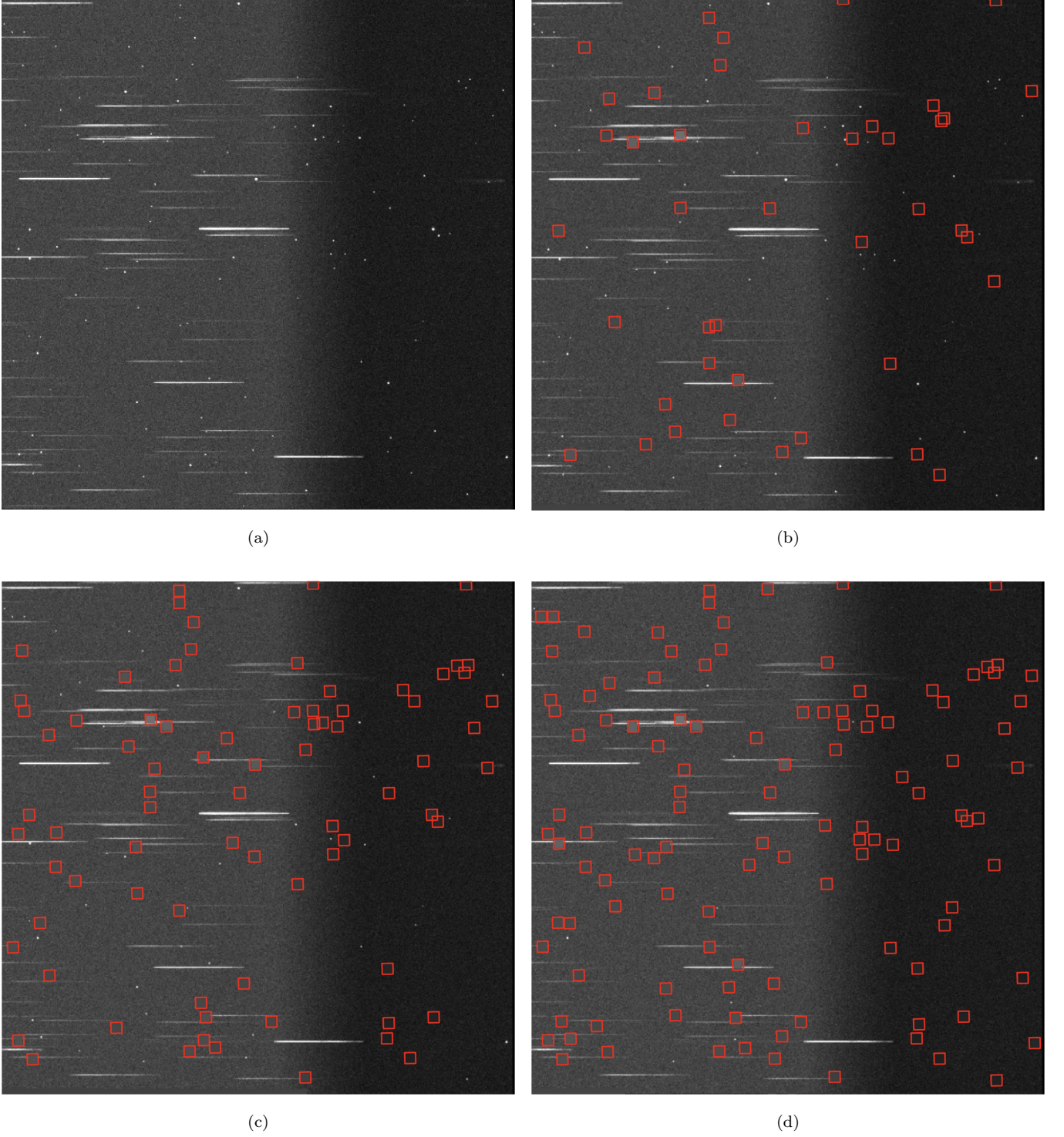


Figure 8. Schematic diagrams of slitless spectral images with varying masked ratios applied to the zero-order image: (a) masked ratio = 0; (b) masked ratio = 0.3; (c) masked ratio = 0.6; (d) masked ratio = 0.9. The masked ratio indicates the fraction of the zero-order image that is obscured, while the spectral lines remain fully visible. The red bounding boxes in the figure indicate the positions of the mask region. There are fragmentation risks where 1D spectra are split by masks, but in real images, the fragmentation risks, as illustrated in (d), do not occur.

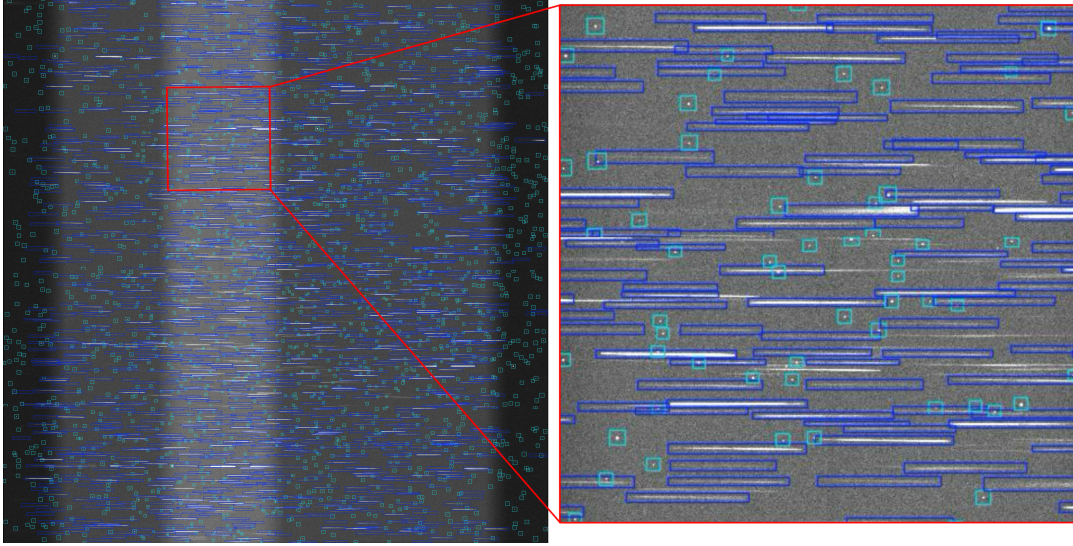


Figure 9. The missed detection in slitless spectroscopic images of medium-density regions within the Galaxy and nearby galaxies.

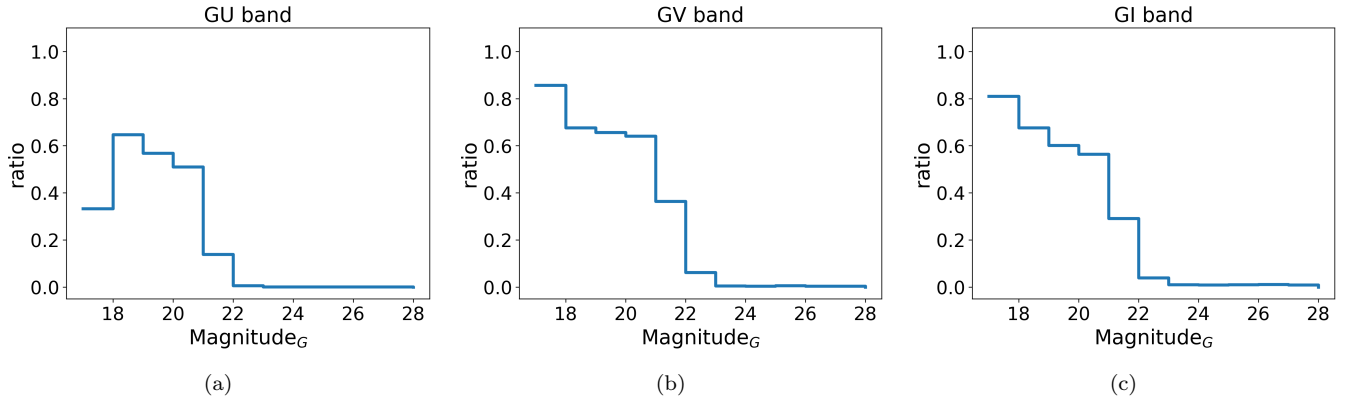


Figure 10. The detection rates of different magnitudes in the GU, GI, and GV bands for galaxy.

where S_* denotes the number of signal electrons; S_S is the average number of sky background electrons per pixel; n_p is the number of pixels used to measure the signal; n_s is the number of pixels used to estimate the background; t represents the exposure time; dc is the dark current; R is the readout noise; G is the gain; and σ_f is the quantization error, modeled as an independent uniform random variable in the range $[-0.5, 0.5]$, galaxies with the same flux as stars typically have lower SNRs due to their extended profiles (i.e., a larger n_p), leading to reduced visibility in the images, as shown in Figure 5 (c). As a result, they are more difficult to detect, leading to lower detection rates at equivalent magnitudes. We separately examined the model's detection rate for galaxies, as shown in Figure 10.

For comparative analysis, we also employed the traditional source detection tool Source Extractor alongside the YOLO model, with the results summarized in Table 7. The algorithm of Source Extractor tends to generate a significant number of false detections, yielding far more sources than present in the input catalog. Source Extractor requires manual parameter tuning, and its detection performance is highly sensitive to these settings, limiting its ability to adapt automatically to diverse datasets. In contrast, deep learning-based approaches, such as YOLO, offer greater robustness and scalability, making them more suitable for large-scale data processing pipelines.

6. CONCLUSION

This paper presents a source detection model for slitless spectral images based on the YOLOv8 algorithm. To enhance the detection of small and densely packed sources, a common challenge in astronomical images, we applied

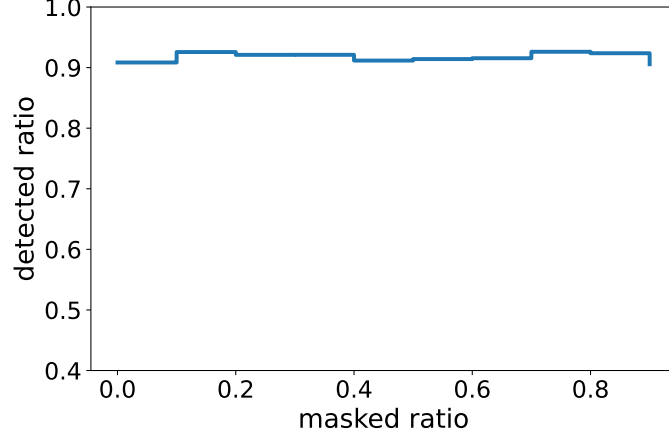


Figure 11. Detection rates of targets brighter than the 22 magnitude at various masked ratios for zero-order images. The masked ratio indicates the proportion of the zero-order image that is covered, with the spectral lines remaining fully visible.

Table 7. The results of Source Extractor.

Region	GU			GV			GI		
	Catalog	SE	YOLO	Catalog	SE	YOLO	Catalog	SE	YOLO
Low	2469	1725	1161	2432	4616	2196	2538	6418	2285
Sculptor	13498	13242	3301	8396	10014	7386	3484	6304	3141
HGL	129	9713	65	223	12309	185	281	14728	212

Note: “Low” represents the low-density region in the Galactic and nearby galaxies, “Sculptor” represents the medium-density region in the Galactic and nearby galaxies, and “HGL” represents the high galactic latitude region. The table displays the number of targets in each waveband and sky region, respectively. “Catalog” represents the number of sources in the catalog, “SE” represents the number detected by the Source Extractor method, and “YOLO” represents the number detected by the method used in this study.

a simple cropping strategy. This process divided 1,560 full-frame images into 24,960 sub-images for training and validation, substantially improving performance on crowded targets. And the main results are summarized as follows:

1. The model achieved precise detection on the validation set, with precision/recall of 88.6%/90.4% for spectral lines and 87.0%/80.8% for zeroth-order images.
2. On the test set, it maintained $>80\%$ detection for sources brighter than 21 mag (medium-density regions) and 20 mag (low-density regions) in the Galactic and nearby galaxy regions, and $>70\%$ for sources brighter than 18 mag in high Galactic latitude regions.
3. Compared to traditional source detection methods, the YOLO approach offers superior performance and a higher level of automation.

The model developed in this study is capable of outputting both the number of sources in the image and their positional coordinates. In future work, we plan to explore the possibility of performing spectral extraction based on these coordinates.

ACKNOWLEDGMENTS

This work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences, Grant No. XDB0550100. J.L is supported by the National Natural Science Foundation of China (NSFC; grant No. 12273027). Man I Lam is supported by the National Natural Science Foundation of China (NSFC; grant No. 12373048).

REFERENCES

- Anthore, A., & Cornu, D. 2024, in EAS2024, European Astronomical Society Annual Meeting, 1350.
<https://ui.adsabs.harvard.edu/abs/2024eas..conf.1350A>
- Bañados, E., Le Brun, V., Belladitta, S., et al. 2025, MNRAS, doi: [10.1093/mnras/staf1274](https://doi.org/10.1093/mnras/staf1274)
- Boyett, K., Mascia, S., Pentericci, L., et al. 2022, ApJL, 940, L52, doi: [10.3847/2041-8213/ac9f17](https://doi.org/10.3847/2041-8213/ac9f17)
- Brammer, G. 2019, Grizli: Grism redshift and line analysis software,, Astrophysics Source Code Library, record ascl:1905.001
<https://ui.adsabs.harvard.edu/abs/2019ascl.soft05001B>
- Carley, E. 2020, in EGU General Assembly Conference Abstracts, EGU General Assembly Conference Abstracts, 5109, doi: [10.5194/egusphere-egu2020-5109](https://doi.org/10.5194/egusphere-egu2020-5109)
- Clowes, R. G., Emerson, D., Smith, M. G., et al. 1980, MNRAS, 193, 415, doi: [10.1093/mnras/193.3.415](https://doi.org/10.1093/mnras/193.3.415)
- Copin, Y., Fumana, M., Mancini, C., et al. 2025, Astronomy & Astrophysics, doi: [10.1051/0004-6361/202554627](https://doi.org/10.1051/0004-6361/202554627)
- Cornu, D., Salomé, P., Semelin, B., et al. 2024, Astronomy & Astrophysics, 690, A211, doi: [10.1051/0004-6361/202449548](https://doi.org/10.1051/0004-6361/202449548)
- Cortes, C., & Vapnik, V. N. 1995, Machine Learning, 20, 273. <https://api.semanticscholar.org/CorpusID:52874011>
- Dalal, N., & Triggs, B. 2005, in 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 1, 886–893 vol. 1, doi: [10.1109/CVPR.2005.177](https://doi.org/10.1109/CVPR.2005.177)
- Doyon, R., Willott, C. J., Hutchings, J. B., et al. 2023, PASP, 135, 098001, doi: [10.1088/1538-3873/acd41b](https://doi.org/10.1088/1538-3873/acd41b)
- Engels, D., Groote, D., Hagen, H. J., & Reimers, D. 1988, in Astronomical Society of the Pacific Conference Series, Vol. 2, Optical Surveys for Quasars, ed. P. Osmer, M. M. Phillips, R. Green, & C. Foltz, 143.
<https://ui.adsabs.harvard.edu/abs/1988ASPC....2..143E>
- Engels, D., Hagen, H. J., Cordis, L., et al. 1998, A&AS, 128, 507, doi: [10.1051/aas:1998390](https://doi.org/10.1051/aas:1998390)
- Euclid Collaboration, Mellier, Y., Abdurro'uf, et al. 2025a, A&A, 697, A1, doi: [10.1051/0004-6361/202450810](https://doi.org/10.1051/0004-6361/202450810)
- Euclid Collaboration, Aussel, H., Tereno, I., et al. 2025b, arXiv e-prints, arXiv:2503.15302, doi: [10.48550/arXiv.2503.15302](https://doi.org/10.48550/arXiv.2503.15302)
- Euclid Collaboration, Jahnke, K., Gillard, W., et al. 2025c, A&A, 697, A3, doi: [10.1051/0004-6361/202450786](https://doi.org/10.1051/0004-6361/202450786)
- Felzenszwalb, P. F., Girshick, R. B., McAllester, D., & Ramanan, D. 2010, IEEE Transactions on Pattern Analysis and Machine Intelligence, 32, 1627, doi: [10.1109/TPAMI.2009.167](https://doi.org/10.1109/TPAMI.2009.167)
- Freund, Y., & Schapire, R. E. 1997, Journal of Computer and System Sciences, 55, 119, doi: <https://doi.org/10.1006/jcss.1997.1504>
- Girshick, R. 2015, in Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), ICCV '15 (USA: IEEE Computer Society), 1440–1448, doi: [10.1109/ICCV.2015.169](https://doi.org/10.1109/ICCV.2015.169)
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. 2014, in 2014 IEEE Conference on Computer Vision and Pattern Recognition, 580–587, doi: [10.1109/CVPR.2014.81](https://doi.org/10.1109/CVPR.2014.81)
- Greene, T. P., Chu, L., Egami, E., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9904, Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave, ed. H. A. MacEwen, G. G. Fazio, M. Lystrup, N. Batalha, N. Siegler, & E. C. Tong, 99040E, doi: [10.1117/12.2231347](https://doi.org/10.1117/12.2231347)
- Grishin, K., Mei, S., Ilic, S., et al. 2025, A&A, 695, A246, doi: [10.1051/0004-6361/202452119](https://doi.org/10.1051/0004-6361/202452119)
- Gullixson, C. A. 1992, in Astronomical Society of the Pacific Conference Series, Vol. 23, Astronomical CCD Observing and Reduction Techniques, ed. S. B. Howell, 130.
<https://ui.adsabs.harvard.edu/abs/1992ASPC...23..130G>
- Hagen, H. J., Groote, D., Engels, D., & Reimers, D. 1995, A&AS, 111, 195.
<https://ui.adsabs.harvard.edu/abs/1995A&AS..111..195H>
- He, K., Gkioxari, G., Dollár, P., & Girshick, R. 2017, in 2017 IEEE International Conference on Computer Vision (ICCV), 2980–2988, doi: [10.1109/ICCV.2017.322](https://doi.org/10.1109/ICCV.2017.322)
- He, Z., Qiu, B., Luo, A.-L., et al. 2021, Monthly Notices of the Royal Astronomical Society, 508, 2039, doi: [10.1093/mnras/stab2243](https://doi.org/10.1093/mnras/stab2243)
- Jiao, L., Zhang, F., Liu, F., et al. 2019, IEEE Access, 7, 128837, doi: [10.1109/ACCESS.2019.2939201](https://doi.org/10.1109/ACCESS.2019.2939201)
- Kümmel, M., Walsh, J. R., Pirzkal, N., Kuntschner, H., & Pasquali, A. 2009, Publications of the Astronomical Society of the Pacific, 121, 59–72, doi: [10.1086/596715](https://doi.org/10.1086/596715)
- Lienhart, R., & Maydt, J. 2002, in Proceedings. International Conference on Image Processing, Vol. 1, I–I, doi: [10.1109/ICIP.2002.1038171](https://doi.org/10.1109/ICIP.2002.1038171)
- Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. 2017, in 2017 IEEE International Conference on Computer Vision (ICCV), 2999–3007, doi: [10.1109/ICCV.2017.324](https://doi.org/10.1109/ICCV.2017.324)

- Liu, W., Anguelov, D., Erhan, D., et al. 2016, in *Computer Vision – ECCV 2016*, ed. B. Leibe, J. Matas, N. Sebe, & M. Welling (Springer International Publishing).
https://link.springer.com/chapter/10.1007/978-3-319-46448-0_2
- Lowe, D. G. 2004, *International Journal of Computer Vision*, 60, 91.
<https://link.springer.com/article/10.1023/B:VISI.0000029664.99615.94>
- Morishita, T., Abramson, L. E., Treu, T., et al. 2019, *ApJ*, 877, 141, doi: [10.3847/1538-4357/ab1d53](https://doi.org/10.3847/1538-4357/ab1d53)
- Mortara, L., & Fowler, A. 1981, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 290, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, ed. J. C. Geary & D. W. Latham, 28, doi: [10.1117/12.965833](https://doi.org/10.1117/12.965833)
- Naufal, A., Koyama, Y., D'Eugenio, C., et al. 2024, *ApJ*, 977, 58, doi: [10.3847/1538-4357/ad8dcf](https://doi.org/10.3847/1538-4357/ad8dcf)
- Oesch, P. A., Brammer, G., van Dokkum, P. G., et al. 2016, *ApJ*, 819, 129, doi: [10.3847/0004-637X/819/2/129](https://doi.org/10.3847/0004-637X/819/2/129)
- Pasquali, A., Pirzkal, N., Larsen, S., Walsh, J., & Kümmel, M. 2006, *Publications of the Astronomical Society of the Pacific*, 118, 270.
<http://www.jstor.org/stable/10.1086/498731>
- Piaulet-Ghorayeb, C., Benneke, B., Radica, M., et al. 2024, *ApJL*, 974, L10, doi: [10.3847/2041-8213/ad6f00](https://doi.org/10.3847/2041-8213/ad6f00)
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. 2016, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788, doi: [10.1109/CVPR.2016.91](https://doi.org/10.1109/CVPR.2016.91)
- Reimers, D., Koehler, T., & Wisotzki, L. 1996, *A&AS*, 115, 235.
<https://ui.adsabs.harvard.edu/abs/1996A&AS..115..235R>
- Reimers, D., & Wisotzki, L. 1997, *The Messenger*, 88, 14.
<https://ui.adsabs.harvard.edu/abs/1997Msngr..88..14R>
- Ren, S., He, K., Girshick, R., & Sun, J. 2017, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39, 1137, doi: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031)
- Ryan, R. E., Casertano, S., & Pirzkal, N. 2018a, *Linear Reconstruction of Grism Spectroscopy I. Simulation and Extraction Examples*, *Instrument Science Report WFC3 2018-13*, 18 pages
<https://ui.adsabs.harvard.edu/abs/2018wfc..rept...13R>
- Ryan, Jr., R. E., Casertano, S., & Pirzkal, N. 2018b, *PASP*, 130, 034501, doi: [10.1088/1538-3873/aaa53e](https://doi.org/10.1088/1538-3873/aaa53e)
- Sekachev, B., Manovich, N., & Zhavoronkov, A. 2019, *Computer Vision Annotation Tool, 0.5.1 Zenodo*, doi: [10.5281/zenodo.3497106](https://doi.org/10.5281/zenodo.3497106)
- Sun, F., Egami, E., Pirzkal, N., et al. 2023, *ApJ*, 953, 53, doi: [10.3847/1538-4357/acd53c](https://doi.org/10.3847/1538-4357/acd53c)
- Varela, L., Boucheron, L., Malone, N., & Spurlock, N. 2019, in *Advanced Maui Optical and Space Surveillance Technologies Conference*, ed. S. Ryan, 89.
<https://ui.adsabs.harvard.edu/abs/2019amos.confE..89V>
- Wang, F., Hu, T., & Gu, M. 2023, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 12511, *Third International Conference on Computer Vision and Data Mining (ICCVDM 2022)*, ed. T. Zhang & T. Yang, 125111C, doi: [10.1117/12.2659975](https://doi.org/10.1117/12.2659975)
- Wei, C., Zhang, X., Liu, D., et al. 2023, *CSST Main Survey Simulation Data: Cycle 6 Data Product Description*, *China Science and Technology Cloud (CSTCloud)*
https://pan.cstcloud.cn/web/view_iframe.html?fid=585816359305371&shareId=zRF7mu3ET5c
- Wei, C., Zhang, X., Liu, D., et al. 2024, *CSST Main Survey Simulation Data: Cycle 9 Data Product Description*, *Kingsoft Docs*
<https://www.kdocs.cn/l/cmIkLt4cfzxw?from=docs>
- Wei, C.-L., Li, G.-L., Fang, Y.-D., et al. 2025, *Research in Astronomy and Astrophysics*, in press
- Wen, R., Zheng, X. Z., Han, Y., et al. 2024, *Monthly Notices of the Royal Astronomical Society*, 528, 2770, doi: [10.1093/mnras/stae157](https://doi.org/10.1093/mnras/stae157)
- Willott, C. J., Doyon, R., Albert, L., et al. 2022, *Publications of the Astronomical Society of the Pacific*, 134, 025002, doi: [10.1088/1538-3873/ac5158](https://doi.org/10.1088/1538-3873/ac5158)
- Wisotzki, L. 1994, in *IAU Symposium*, Vol. 161, *Astronomy from Wide-Field Imaging*, ed. H. T. MacGillivray, 723.
<https://ui.adsabs.harvard.edu/abs/1994IAUS..161..723W>
- Wisotzki, L., Koehler, T., Groote, D., & Reimers, D. 1996, *A&AS*, 115, 227. <https://ui.adsabs.harvard.edu/abs/1996A&AS..115..227W>
- Worseck, G., Wisotzki, L., & Selman, F. 2008, *Astronomy & Astrophysics*, 487, 539–554, doi: [10.1051/0004-6361/200810157](https://doi.org/10.1051/0004-6361/200810157)
- Xing, Y., Yi, Z., Liang, Z., et al. 2023, *The Astrophysical Journal Supplement Series*, 269, 59, doi: [10.3847/1538-4365/ad0551](https://doi.org/10.3847/1538-4365/ad0551)
- Zhan, H. 2021, *Chinese Science Bulletin*, 66, 1290, doi: [10.1360/TB-2021-0016](https://doi.org/10.1360/TB-2021-0016)
- Zhang, X., Fang, Y.-D., Wei, C.-L., et al. 2025, *Research in Astronomy and Astrophysics*, in press
- Zhao, Q., Sheng, T., Wang, Y., et al. 2019, in *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence*, AAAI'19/IAAI'19/EAAI'19 (AAAI Press), doi: [10.1609/aaai.v33i01.33019259](https://doi.org/10.1609/aaai.v33i01.33019259)

Zhao, Z.-Q., Zheng, P., Xu, S.-T., & Wu, X. 2019, IEEE Transactions on Neural Networks and Learning Systems, 30, 3212, doi: [10.1109/TNNLS.2018.2876865](https://doi.org/10.1109/TNNLS.2018.2876865)

Zhou, Y., Liu, C., Tian, H., Zhang, X., & Li, N. 2025, The Astronomical Journal, 170, 256, doi: [10.3847/1538-3881/adee1c](https://doi.org/10.3847/1538-3881/adee1c)