



Accurately Estimating Redshifts from CSST Slitless Spectroscopic Survey Using Deep Learning

Xingchen Zhou^{1,2}, Yan Gong^{1,2,3} , Xin Zhang^{1,2}, Nan Li^{1,2}, Xian-Min Meng^{1,2}, Xuelei Chen^{1,2,4} , Run Wen^{5,6} , Yunkun Han⁷ , Hu Zou¹ , Xian Zhong Zheng^{5,6} , Xiaohu Yang^{8,9} , Hong Guo¹⁰ , and Pengjie Zhang^{8,9}

¹ National Astronomical Observatories, Chinese Academy of Sciences, 20A Datun Road, Beijing 100101, People's Republic of China; gongyan@bao.ac.cn
² Science Center for China Space Station Telescope, National Astronomical Observatories, Chinese Academy of Sciences, 20A Datun Road, Beijing 100101, People's Republic of China

³ University of Chinese Academy of Sciences, Beijing, 100049, People's Republic of China

⁴ Centre for High Energy Physics, Peking University, Beijing 100871, People's Republic of China

⁵ Purple Mountain Observatory, Chinese Academy of Sciences, 10 Yuanhua Road, Nanjing 210023, People's Republic of China

⁶ School of Astronomy and Space Sciences, University of Science and Technology of China, Hefei 230026, People's Republic of China

⁷ Yunnan Observatories, Chinese Academy of Sciences, 396 Yangfangwang, Guandu District, Kunming 650216, People's Republic of China

⁸ Department of Astronomy, School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China

⁹ Tsung-Dao Lee Institute and Key Laboratory for Particle Physics, Astrophysics and Cosmology, Ministry of Education, Shanghai 201210, People's Republic of China

¹⁰ Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, People's Republic of China

Received 2024 July 18; revised 2024 October 17; accepted 2024 October 24; published 2024 December 3

Abstract

Chinese Space Station Telescope (CSST) has the capability to conduct a slitless spectroscopic survey simultaneously with a photometric survey. The spectroscopic survey will measure slitless spectra, potentially providing more accurate estimations of galaxy properties, particularly redshifts, compared to using broadband photometry. CSST relies on these accurate redshifts to use baryon acoustic oscillations (BAOs) and other probes to constrain the cosmological parameters. However, due to the low resolution and signal-to-noise ratio of slitless spectra, measurement of redshifts is significantly challenging. In this study, we employ a Bayesian neural network (BNN) to assess the accuracy of redshift estimations from slitless spectra anticipated to be observed by CSST. The simulation of slitless spectra is based on real observational data from the early data release of the Dark Energy Spectroscopic Instrument (DESI-EDR) and the 16th data release of the Baryon Oscillation Spectroscopic Survey (BOSS-DR16), combined with the 9th data release of the DESI Legacy Survey (DESI LS DR9). The BNN is constructed employing a transfer learning technique, by appending two Bayesian layers after a convolutional neural network, leveraging the features learned from the slitless spectra and corresponding redshifts. Our network can provide redshift estimates along with corresponding uncertainties, achieving an accuracy of $\sigma_{\text{NMAD}} = 0.00063$, outlier percentage $\eta = 0.92\%$, and weighted mean uncertainty $\bar{E} = 0.00228$. These results successfully fulfill the requirement of $\sigma_{\text{NMAD}} < 0.005$ for BAO and other studies employing CSST slitless spectroscopic surveys.

Unified Astronomy Thesaurus concepts: Galaxies (573); Cosmology (343)

1. Introduction

Redshifts are one of the fundamental quantities for studying galaxies. The most accurate redshifts are determined through observing and analyzing high-resolution spectra from galaxies. However, obtaining high-resolution spectra is a time-consuming task, especially for high-redshift and faint sources, which require hours of observation to successfully measure their redshifts. As a result, photometric redshifts, estimated from several photometric measurements, become a necessary option for most sources observed in ongoing and future cosmological surveys. However, the best accuracy of σ_{NMAD} can barely achieve ~ 0.01 among photometric redshift estimation endeavors utilizing real observational data (H. Zou et al. 2019; S. Schuldt et al. 2021; E. Jones et al. 2024; M. Treier et al. 2024). This accuracy severely hinders certain cosmological studies using techniques such as baryon acoustic oscillations (BAOs; B. Bassett & R. Hlozek 2010) and redshift-space distortions (RSDs; A. J. S. Hamilton 1998). To match the

accuracy required by these cosmological studies and the survey speed of current photometric surveys, a compromise solution exists: slitless spectra. Slitless spectra represent a category of low-resolution spectroscopy performed without a narrow slit, which typically allows only light from a small region to be diffracted. Current and future photometric surveys, such as the Euclid Space Telescope (Euclid; Euclid Collaboration et al. 2024), James Webb Space Telescope (JWST; P. A. Sabelhaus & J. E. Decker 2004), Nancy Grace Roman Space Telescope (Roman),¹¹ and Chinese Space Station Telescope (CSST; H. Zhan 2018; Y. Gong et al. 2019), all include modules to observe slitless spectra for galaxies.

CSST is a 2 m space telescope designed for photometric surveys across seven bands, ranging from near-ultraviolet to near-infrared. The slitless spectroscopy module, which includes three bands (*GU*, *GV*, and *GI*), operates alongside the photometric module, enabling simultaneous photometric and slitless spectroscopic observations. These three bands can reach a 5σ magnitude limit of 23.2, 23.4, and 23.2 for point sources, respectively, with a low spectroscopic resolution of each band of $R = \lambda/\Delta\lambda \geq 200$ (Y. Gong et al. 2019). For extended



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

¹¹ <https://roman.gsfc.nasa.gov/>

sources such as galaxies, slitless spectra can be significantly affected by observational and instrumental effects, challenging the one-dimensional spectrum extraction procedure and thus resulting in low-resolution and signal-to-noise-ratio (SNR) spectra. These challenges cause significant difficulties in the recognition of emission and absorption lines, breaks, and other spectroscopic features. As a result, galaxy properties such as redshift and line fluxes estimated from these spectra can be highly inaccurate, leading to low accuracy even comparable to that derived from broadband photometry. Addressing the challenge of successfully measuring these galaxy properties from such low-resolution and SNR slitless spectra remains an urgent problem.

Machine Learning (ML), particularly Deep Learning (DL) algorithms (also known as neural networks), offers a potential solution to the challenges. This algorithm can effectively learn the inherent correlations between inputs and outputs using large data sets, making them well suited to handle data significantly affected by instrumental or other forms of noise. In the astronomical and cosmological communities, neural networks have gained prominence in recent years, achieving applications across various fields. The multilayer perceptron (MLP), a simple neural network, has been applied to estimate photometric redshifts from multiband photometric measurements (A. A. Collister & O. Lahav 2004; I. Sadeh 2014; X. Zhou et al. 2022a), surpassing the accuracy achieved by traditional spectral energy distribution (SED) fitting methods.

Furthermore, state-of-the-art convolutional neural networks (CNNs; Y. Lecun et al. 1998), which excel in directly processing images, have become indispensable in astronomical and cosmological analysis. Applications include deriving photometric redshifts or other quantities from galaxy images (J. Pasquet et al. 2019; M. Tewes et al. 2019; B. Henghes et al. 2022; X. Zhou et al. 2022a; Z. Zhang et al. 2024), discovering strong lensing systems or mergers (C. Schaefer et al. 2018; W. J. Pearson et al. 2019; Z. He et al. 2020; R. Li et al. 2020; S. Rezaei et al. 2022; A. R. Arendt et al. 2024), and constraining cosmological parameters from large-scale structures or weak gravitational lensing (A. Gupta et al. 2018; S. Pan et al. 2020; J. Fluri et al. 2022; H. J. Hortúa et al. 2023; Z. Min et al. 2024).

In addition to processing two-dimensional arrays, CNNs can be adapted to handle one-dimensional sequences or three-dimensional data cubes. Spectra are one-dimensional sequences containing redshift or other information, which can be effectively extracted by 1D CNNs. The application of 1D CNNs for deriving redshifts from spectra is extensively researched (N. Busca & C. Balland 2018; F. Rastegarnia et al. 2022).

Unlike traditional fitting methods that produce both redshifts and uncertainties, deep learning methods typically provide only redshift values. Recognizing the importance of uncertainties in cosmological studies, Bayesian neural networks (BNNs; D. J. C. MacKay 1995; C. Blundell et al. 2015; Y. Gal & Z. Ghahramani 2015), which can output both point estimations and uncertainties, have gained significant attention, especially in cosmological studies (L. Perreault Levasseur et al. 2017; H. J. Hortúa et al. 2020; X. Zhou et al. 2022b; H. J. Hortúa et al. 2023; E. Jones et al. 2024). By assigning probability distributions to each weight in the network, BNNs can capture and propagate uncertainties from data and the neural network

itself to the output, providing not only point predictions but also confidence intervals or posterior distributions.

Although deep learning algorithms offer advantages in providing better accuracy, higher speed, and direct processing of raw data, several challenges need careful consideration. Since deep learning models heavily rely on training data, obtaining abundant and representative data for observation is a primary concern. Specifically, for redshift estimation, a large data set with high-quality spectroscopic redshifts is essential. Fortunately, several ongoing and planned spectroscopic surveys, such as the Dark Energy Spectroscopic Instrument (DESI; DESI Collaboration et al. 2016), Prime Focus Spectrograph (PFS; N. Tamura et al. 2016), MULTiplexed Spectroscopic Telescope (MUST),¹² MegaMapper (D. J. Schlegel et al. 2022), and Wide-field Spectroscopic Telescope (WST; V. Mainieri et al. 2024), aim to obtain a substantial number of galaxy spectra with accurate redshifts. Including completed surveys like zCOSMOS (S. J. Lilly et al. 2007), VIMOS-VLT Deep Survey (VVDS; O. Le Fevre et al. 2013), Sloan Digital Sky Survey (SDSS; R. Ahumada et al. 2020), and Baryon Oscillation Spectroscopic Survey (BOSS; K. S. Dawson et al. 2013), a sufficient and representative training set for redshift estimation can be achieved.

In this work, we employ a deep learning technique to estimate spectroscopic redshifts for slitless spectra expected to be observed by CSST. The slitless spectra are simulated based on real spectroscopic observations. Considering the redshift coverage and survey fields, we employ data from the DESI early data release (DESI-EDR; DESI Collaboration 2023) and BOSS 16th data release (BOSS-DR16; K. S. Dawson et al. 2013), combined with the 9th data release of DESI Legacy Surveys (DESI LS DR9; D. Schlegel et al. 2021). DESI-EDR has made available 1.2 million high-resolution spectra of galaxies and quasars collected during the Survey Validation (SV) phase for target selections. Since the number of sources in DESI-EDR is limited, we supplement our slitless spectrum data set with BOSS data, which shares a similar pipeline for the measurement of spectroscopic redshifts, to increase the data size for training the neural network model. After obtaining the slitless spectra, we train a 1D BNN with these spectra and their corresponding accurate spectroscopic redshifts, and analyze the accuracy that can be achieved for the CSST slitless spectroscopic survey.

The structure of this paper is organized as follows: we briefly describe CSST slitless spectra simulation software and then explain the generation of the mock slitless spectra in Section 2. The neural network methods including CNNs and BNNs are introduced in Section 3. Then, we demonstrate our results in Section 4. The limitations of the current study are extensively discussed in Section 5. Finally, this paper is concluded in Section 6.

2. Mock Data

In this section, we first introduce the slitless spectra simulation software in the CSST data analysis pipeline, and then explain the data generation procedure of slitless spectra using this software from real spectroscopic observations.

¹² <https://must.astro.tsinghua.edu.cn/en>

2.1. Slitless Spectra Simulation Software

The simulation software for slitless spectra is an integral part of the CSST data analysis pipeline, with the code available online.¹³ We provide a brief overview of the workflow here, and interested readers are recommended to consult X. Zhang et al. (2024, in preparation) for detailed information. This software utilizes SEDs and morphological parameters of galaxies to generate mock spectra. Initially, the dispersion curve for the grating is determined through a fitting process that considers the spectroscopic properties of the CSST's slitless spectrum. Following this, the energy profile of the galaxy, assumed to be a Sérsic profile (J. L. Sérsic 1963, 1968) derived from its morphological parameters, is converted into a pixelated galaxy image. Each pixel of the galaxy image undergoes dispersion based on the dispersion curve specific to the CSST grating, in conjunction with the sensitivity curve of the CSST instrument and the galaxy's SED. Finally, all dispersed components are integrated into a two-dimensional slitless spectral image. Additionally, instrumental effects are simulated using a point-spread function (PSF), assumed to be a 2D Gaussian distribution with a full width at half-maximum (FWHM) of 0''.3. The sky backgrounds, including zodiacal and earthshine components, are computed as 0.019, 0.214, and $0.329 \text{ e}^{-\text{s}^{-1}} \text{ pixel}^{-1}$ for the *GU*, *GV*, and *GI* bands, respectively. To mitigate the effects of instrumental and background noise, we coadd spectra from four exposures, each with a duration of 150 s. Following these procedures, we generate first-order spectral images expected to be observed by CSST, from which the 1D spectra and the corresponding errors can be extracted.

2.2. Data Generation

To realistically simulate our slitless spectra, we utilize spectroscopic observations from the Dark Energy Spectroscopic Instrument (DESI) and the Baryon Oscillation Spectroscopic Survey (BOSS). DESI is an ongoing spectroscopic survey conducted on the Mayall 4 m telescope at Kitt Peak National Observatory. Over its 5 yr mission, DESI aims to observe spectra for more than 30 million galaxies and quasars across 14,000 square degrees of sky (DESI Collaboration et al. 2016). Recently, DESI has released its Early Data Release (EDR), which includes spectroscopic data for 1.8 million targets observed during the SV phase conducted from 2020 December to 2021 June (DESI Collaboration 2023).

We select sources from the EDR spectroscopic redshift catalog using the following criteria:

```
SV_PRIMARY == True
MASKBITS == 0
SPECTYPE == GALAXY
ZWARN == 0
FLUX_G, R, Z > 0
FLUX_IVAR_G, R, Z > 0
MORPHTYPE != PSF
```

(1)

Here, `SV_PRIMARY` indicates the best recommended redshift if the same source appears multiple times in the catalog, while `MASKBITS` is the bit-wise mask indicating that the source touches a pixel in a masked area. `SPECTYPE` and `ZWARN` are

¹³ https://csst-tb.bao.ac.cn/code/zhangxin/sls_1d_spec

source classification and indicators of potential issues in the spectroscopic redshift measured by Redrock, a commonly used redshift-fitting software.¹⁴ We further control the quality of sources by applying constraints on photometric measurements in the *g*, *r*, and *z* bands of the DESI legacy imaging survey (A. Dey et al. 2019). `MORPHTYPE` indicates the Tractor model used to fit the source during photometric measurement. This constraint ensures that the selected sources are extended, allowing for accurate morphological parameter measurements. It should be noted that some PSF sources are spectroscopically classified as galaxies. These PSF models are assigned probably due to the resolution of $\sim 1''.0$ in imaging data of DESI legacy survey, and we simply exclude these galaxies in our data set. To obtain the morphological parameters required to derive our slitless spectra, we match the selected sources with the sweep catalog of the DESI legacy survey DR9¹⁵ and retrieve the morphological parameters, including effective radius r_{eff} , Sérsic index n , and two ellipticity components ϵ_1 , ϵ_2 and their variance. And we perform another selection to filter the sources with valid morphological measurements:

$$\begin{aligned} \text{SHAPE_R} &> 0 \\ \text{SHAPE_IVAR_R} &> 0 \\ \text{SHAPE_E1_IVAR} &> 0 \\ \text{SHAPE_E2_IVAR} &> 0 \\ \text{SERSIC_IVAR} &> 0 \end{aligned} \quad (2)$$

and then we calculate the axis ratio b/a and position angle ϕ using the following equations as recommended by DESI:¹⁶

$$\begin{aligned} |\epsilon| &= \sqrt{\epsilon_1^2 + \epsilon_2^2}, \\ \frac{b}{a} &= \frac{1 - |\epsilon|}{1 + |\epsilon|}, \\ \phi &= \frac{1}{2} \arctan \frac{\epsilon_2}{\epsilon_1}. \end{aligned} \quad (3)$$

This selection process results in approximately 180,000 sources with high-quality spectroscopic redshifts, with the distribution illustrated in the blue histogram in Figure 1. The majority of these sources are bright galaxy samples (BGSs) and luminous redshift galaxies (LRGs) in DESI, since these categories of sources are larger in size, and consequently, yield measurable morphological parameters. It is acknowledged that the morphological parameters of certain galaxies exhibit significant uncertainties. Furthermore, the morphological variations of some galaxies in spatial dimensions and across different spectral bands cannot be fully captured by single Sérsic profiles, thereby diminishing the realness of simulated slitless spectra. These effects are beyond the scope of this study and will be addressed in future research.

After obtaining the redshifts and morphological parameters, the next step involves acquiring the SED for each source to simulate slitless spectra. The spectroscopic redshifts of the sources have been determined using model spectra fitted by Redrock, and all redshift warning flags are zero, indicating no

¹⁴ <https://github.com/desihub/redrock>

¹⁵ <https://www.legacysurvey.org/dr9/files/#sweep-catalogs-region-sweep>

¹⁶ <https://www.legacysurvey.org/dr10/catalogs/#ellipticities>

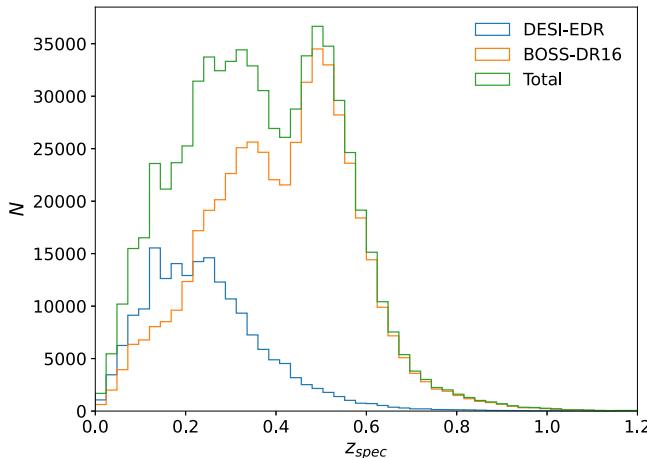


Figure 1. Spectroscopic redshift distributions for selected sources in DESI-EDR, BOSS-DR16, and total.

issues with the fitting process. This allows us to use the model spectra to accurately represent the SED of each source. These model spectra can be constructed using the COEFF provided in the DESI-EDR catalog, combining Redrock templates, or accessed via the SPectra Analysis & Retrievable Catalog Lab (SPARCL).¹⁷ The model spectra obtained through both approaches are the same, and we choose the latter.

The 180,000 sources selected from DESI-EDR are insufficient for investigating the redshift accuracy in the CSST slitless spectroscopic survey, since the redshift distribution displayed in Figure 1 can barely reach 0.8, much lower than the estimated redshift limit of about 1.5 expected to be observed by CSST (Y. Gong et al. 2019). Furthermore, the estimations at high redshift can be problematic without enough training data. To address these challenges, we supplement our slitless spectra utilizing data from the BOSS. BOSS is a spectroscopic survey primarily targeting luminous red galaxies (LRGs) up to $z \sim 0.7$ and quasars (QSOs) at redshifts $2.2 < z < 3$, aimed at detecting the characteristic scale imprinted by BAOs in the early Universe. Over its 5 yr observation period, BOSS has measured spectra for approximately 4 million sources, covering 10,000 square degrees (K. S. Dawson et al. 2013). For this work, we use 16th data release from BOSS (BOSS-DR16).

Similar to our approach with DESI data, we select galaxies with a spectroscopic redshift warning ZWARN==0, produced by the Redrock software, and match these sources with the DESI LS DR9. We exclude sources modeled as a PSF and without reasonable photometric measurements in the g , r , and z bands and those lacking valid morphological parameters. This results in a selection of 450,000 galaxies, for which we download their model spectra using SPARCL. Finally, we merge the sources from DESI and BOSS, obtaining approximately 600,000 sources in total.

The spectroscopic redshift distributions are illustrated in Figure 1. We notice that most sources from DESI-EDR are at low redshifts, while high-redshift sources up to $z \sim 1$ are supplemented by BOSS-DR16. The distributions of morphological parameters, including effective radius r_{eff} , Sérsic index n , axis ratio b/a , and position angle ϕ are illustrated in Figure 2. Notably, galaxies with a Sérsic index of 6 dominate,

particularly those from BOSS-DR16. This selection bias is as expected, as the Sérsic index positively correlates with galaxy size and luminosity. Hence, for valid morphological measurements and accurate spectroscopic redshift fitting, the sources are expected to be larger in size and brighter in luminosity, with large Sérsic indices.

After obtaining the model spectra and morphological parameters for sources in DESI-EDR and BOSS-DR16, we simulate the slitless spectra using the simulation software mentioned in Section 2.1. The signal-to-noise ratio (SNR) of the simulated CSST slitless spectra in GU , GV , GI , and total are illustrated in Figure 3. We notice that the SNRs of GI are the best, reaching a peak at about 2, while most SNRs of GU and GV bands are lower than 1, especially for the GU band. The significantly low SNR in the GU band can be attributed to the fact that the wavelength coverage of the DESI and BOSS spectrographs does not fully encompass the wavelength coverage of the GU band of CSST. Consequently, there are negative data points in the model spectra over the GU band, and thus the slitless spectra in the GU band are predominantly characterized by Gaussian noise. Additionally, the total SNRs of these spectra peak at ~ 1 , indicating the signal and noise are at a similar level. In Figure 4, we display two examples of simulated first-order slitless spectral images in GU , GV , and GI bands, and the corresponding extracted one-dimensional spectra are shown in Figure 5. The SEDs used in simulation are also illustrated, and they are consistent with the spectra. Additionally, the source information including coordinates (R.A. and decl.), spectroscopic redshifts, morphological parameters used in simulation, and SNR in GU , GV , and GI bands are also shown. For the low-redshift source in the left panel, we can clearly recognize the dispersed 2D spectra in the GV and GI bands, and SNRs of the extracted 1D spectra in these two bands are relatively high. For the high-redshift source in the right panel, only a faint 2D spectrum in the GI band can be recognized, with the other two bands dominated by noise; hence, the extracted 1D spectra correspondingly have low SNR. Overall, the slitless spectra are severely affected by background and instrumental noise, and the recognition of spectroscopic features such as breaks and absorption and emission lines is difficult, leading to challenges for successful redshift determinations using traditional approaches such as spectrum fitting or feature identification.

3. Methodology

We employ a CNN to estimate redshifts from slitless spectra expected to be observed by CSST. To satisfy the requirement of some cosmological studies, we further construct a BNN to derive redshift values along with their uncertainties. Our networks are constructed using Keras¹⁸ backend on TensorFlow¹⁹ and TensorFlow-Probability,²⁰ and the relevant codes are publicly available at Github.²¹

3.1. Neural Networks

Since spectra are one-dimensional sequences, we employ a 1D CNN to process them. A CNN is a powerful deep learning model that can learn the internal connections between data and

¹⁸ <https://keras.io/>

¹⁹ <https://www.tensorflow.org/>

²⁰ <https://www.tensorflow.org/probability>

²¹ https://github.com/xczhou-astro/CSST_slitless_spectra

¹⁷ <https://astrosparcl.datelab.noirlab.edu/>

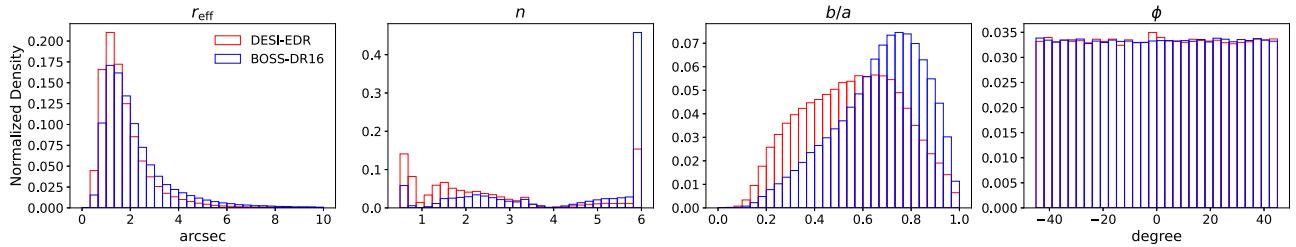


Figure 2. The distribution of four morphological parameters: effective radius r_{eff} , Sérsic index n , axis ratio b/a , and position angle ϕ of sources from DESI-EDR and BOSS-DR16.

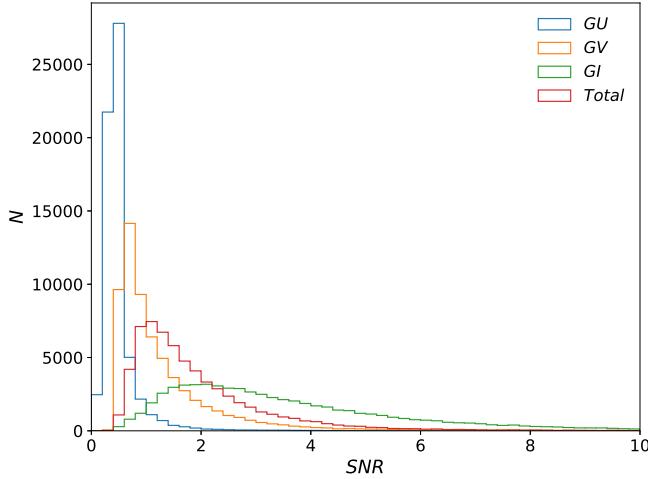


Figure 3. The SNR distributions of simulated slitless spectra in GU , GV , and GI bands and over the whole wavelength range.

labels. Therefore, we expect that our 1D CNN can learn the mapping between slitless spectra and redshifts. To improve its learning ability, we increase the depth of our 1D CNN using ResNet blocks (K. He et al. 2015). This block can effectively reduce the vanishing gradient problem that commonly happens in deep neural networks through skip connections, as illustrated in the right panel of Figure 6. The convolutional layer in the skip connection is applied when this block processes and downsamples the features at the same time. Following X. Zhou et al. (2021), the input to our CNN includes spectra and corresponding errors as a two-channel sequences. Then, the inputs are processed by one convolutional layer with 32 kernels with kernel sizes of 7, followed by a max-pooling to reduce the feature dimension. After this shallow feature extraction layer, we structure eight ResNet blocks to obtain useful features from spectra and further reduce the feature dimension. The 1D convolutional layers in these blocks are all followed by BatchNormalization layers (S. Ioffe & C. Szegedy 2015) to reduce overfitting. And ReLU activation functions (A. F. Agarap 2018) are structured to apply nonlinearity. Subsequently, the features are vectorized to one dimension using global average pooling, and then a dropout layer (N. Srivastava et al. 2014) with drop rate of 0.2 is applied. This dropout layer also functions as overfitting reduction. Finally, the output layer with one neuron is structured. The illustration of the architecture is displayed in the left panel of Figure 6.

The neural network discussed so far can only output redshift values. To output redshifts along with their

uncertainties, we construct a BNN. The uncertainties can be categorized into two types: epistemic uncertainty, which arises from neural network models and can be reduced by averaging ensemble networks with different configurations, and aleatoric uncertainty, which originates from intrinsic corruption of the data and cannot be reduced (A. D. Kiureghian & O. Ditlevsen 2009; E. Hüllermeier & W. Waegeman 2019). A BNN can capture both uncertainties by varying the network configurations and utilizing a specific distribution as the output, with the Gaussian distribution being the most common choice. For more detailed discussion on this network, please refer to H. J. Hortúa et al. (2020) and X. Zhou et al. (2022b). Our BNN is built using a transfer learning technique to transfer the feature extraction part before the final output layer of the CNN and then append two Bayesian layers. The weights from the transferred network are set as fixed, leveraging the features that are tailored to derive the redshift. For the Bayesian layers, three common structures are commonly employed, i.e., Monte Carlo dropout (MC-dropout; Y. Gal & Z. Ghahramani 2015), flipout (Y. Wen et al. 2018), and Multiplicative Normalizing Flows (MNF; C. Louizos & M. Welling 2017) layers. Among these three categories, MC-dropout uses dropout to simulate varying configurations of the network, while the weights in flipout and MNF layers are represented by distributions. Particularly, MNF employs more complicated distributions transformed from a Gaussian distribution using normalizing flows (D. Jimenez Rezende & S. Mohamed 2015). As recommended by another work (X. Zhou et al. 2024, in preparation), we adopt MNFDense layers,²² adapting 50 layers for masked RealNVP normalizing flow (L. Dinh et al. 2016). Since a Gaussian distribution is a reasonable assumption for redshift estimation, the network finally outputs two values, from which a Gaussian distribution can be derived. The illustration of the architecture is also displayed in the left panel of Figure 6.

Bayesian networks inherently introduce uncertainties that are either overestimated or underestimated, deviating from the statistical principle that the coverage probabilities of the sample with the true value within specific confidence intervals correspond to the corresponding confidence intervals (L. Perrault Levasseur et al. 2017; H. J. Hortúa et al. 2020). To assess the calibration status, we can employ the reliability diagram, which plots the coverage probability against the confidence interval. The uncertainties are well calibrated when this diagram exhibits a straight diagonal line. Conversely, calibration is essential before reporting the results. Extensive discussions on calibration techniques can be found in C. Guo et al. (2017) and H. J. Hortúa et al. (2020). In this study, we

²² <https://github.com/janosh/tf-mnf>

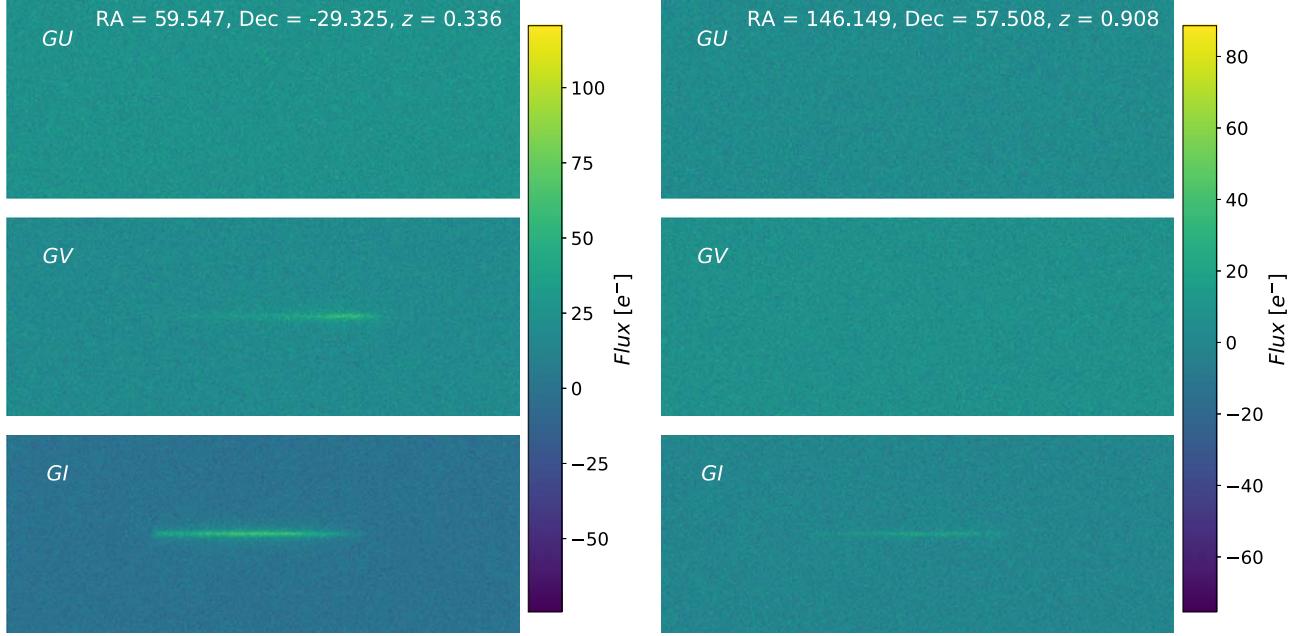


Figure 4. Two examples of simulated first-order slitless spectral images in GU , GV , and GI bands. The coordinates and spectroscopic redshifts are also shown. The corresponding extracted one-dimensional spectra and more information for the two sources are displayed in Figure 5.

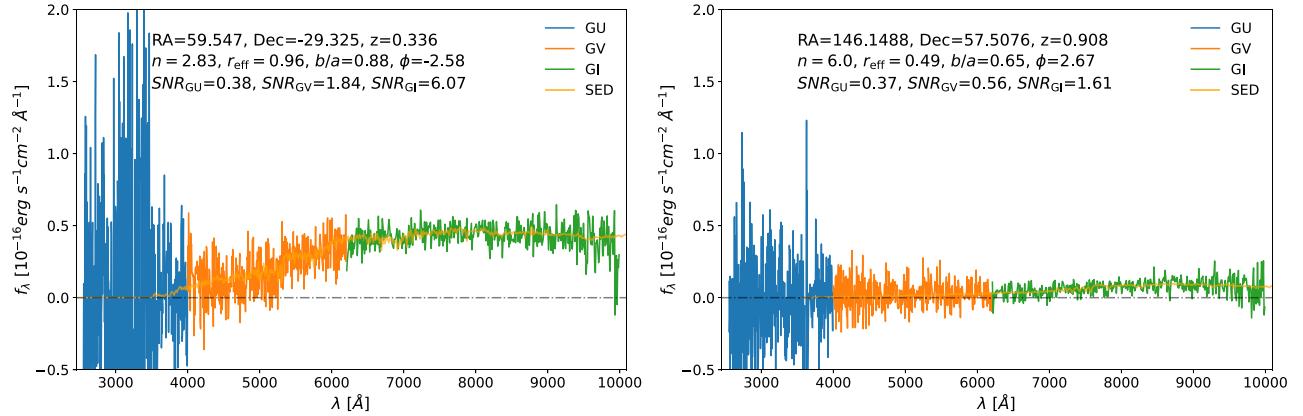


Figure 5. The corresponding one-dimensional spectra extracted from spectral images of sources in Figure 4. The SEDs used in simulation are also illustrated. The black dashed-dotted line indicates zero fluxes. Additionally, the source information including coordinates, spectroscopic redshifts, morphological parameters, and SNR of each band are also shown.

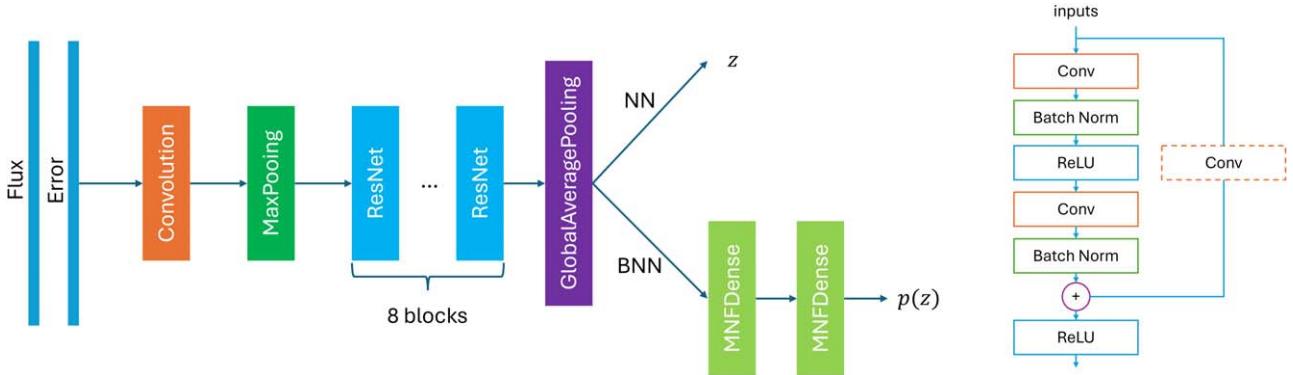


Figure 6. Left: the architecture of the 1D CNN and BNN built upon ResNet blocks. Right: the structure of the ResNet block.

adopt a straightforward after-training calibration technique, Beta calibration (M. Kull et al. 2017), to calibrate the uncertainties. This calibration procedure ensures that the

uncertainties more accurately reflect the true confidence intervals, thereby enhancing the reliability of our redshift estimations.

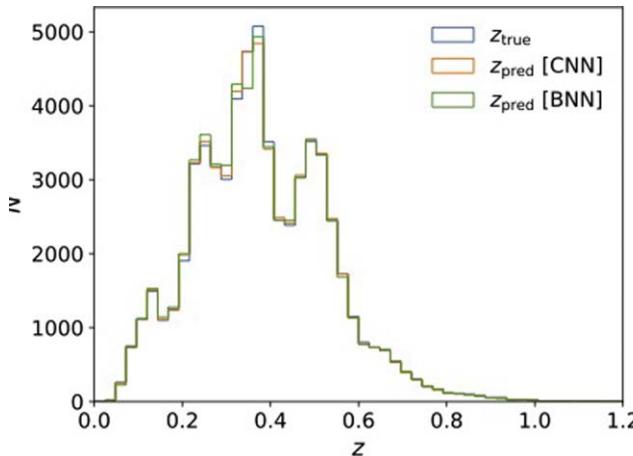


Figure 7. The distributions of true and predicted redshifts from the CNN and BNN for testing data.

3.2. Training

We only consider the spectra in *GV* and *GI* bands, since the SNRs in the *GU* band are significantly lower and the spectra are dominated by noise, as explained in Section 2.2. Before training, our data are split into training, validation, and testing sets as a ratio of 8:1:1.²³ The number of testing set is approximately 60,000, and they are selected based on the expected redshift distribution of the CSST slitless spectroscopic survey. The redshift distribution of the testing set is displayed in Figure 7, and is consistent with the one shown in Y. Gong et al. (2019) in the low-redshift region. To improve the performance of the neural network, we follow the methodology in X. Zhou et al. (2021) to increase the size of the training set by involving their Gaussian realization counterparts created through fluctuating the spectra based on their corresponding errors. This data augmentation technique can effectively amplify the adaptability of the network to the large noise in low-SNR slitless spectra. Here, we use 50 random realizations. For the 1D CNN, we set the loss function and optimizer as logCosh and Adam. LogCosh resembles a traditional mean-absolute-error function but with a differential behavior around 0, and Adam is a stochastic gradient descent optimization method based on the adaptive estimate of first-order and second-order moments (D. P. Kingma & J. Ba 2014). This network is trained for 100 epochs with a batch size of 1024 considering the graphic memory of the GPU. We select the best model with the lowest loss value as our final CNN model and the backbone for the BNN, since the performance of the CNN backbone substantially influences the outcomes of the BNN.

For the BNN, only the weights of the appended MNF layers are optimized in training. This training strategy can leverage the features optimized for estimating redshifts and substantially increase the training speed. The loss function of the BNN is set to be negative log-likelihood (NLL), different from the one employed in the CNN, since BNN outputs a distribution considering both point values and their uncertainties. Note that the labels are solely redshift values, since the uncertainties are naturally derived during the decrease of the loss function. Similarly, we adopt the Adam optimizer and save the model with the lowest loss value. Different from the CNN in post processing, we feed the testing spectra to the BNN for 200

²³ Data are available at <https://pan.cstcloud.cn/s/E6FrFGa6TJA>.

Table 1

Main Hyperparameters that We Tune and Adopt for the CNN and BNN

| | CNN | BNN |
|---------------------------|----------------------|------------------|
| Kernel size of 1D CNN | 7 | ... ^a |
| Dropout rate | 0.2 | ... |
| Number of ResNet blocks | 8 | ... |
| Number of Bayesian layers | ... | 2 |
| Learning rate for Adam | 1e-4 | 1e-4 |
| Loss function | LogCosh ^b | NLL |
| Batch size | 1024 | 1024 |
| Number of augments | 50 | 50 |

Notes.

^a Inherited from CNN through transfer learning.

^b LogCosh and HuberLoss are experimented.

times. Based on these outputs, we calculate the final redshift values and their corresponding uncertainties that account for both epistemic and aleatoric ones.

The main hyperparameters for the CNN and BNN that we tune and adopt are summarized in Table 1. Note that the hyperparameters do not severely affect the performance for both networks as long as the network is deep enough and the training data are sufficient.

4. Results

We employ two metrics to evaluate the performance of the CNN: outlier percentage η and normalized median absolute deviation σ_{NMAD} , defined as follows:

$$\eta = \frac{N_{\Delta z/(1+z_{\text{true}})>0.02}}{N_{\text{total}}}, \quad (4)$$

$$\sigma_{\text{NMAD}} = 1.48 \times \text{median} \left(\left| \frac{\Delta z - \text{median}(\Delta z)}{1 + z_{\text{true}}} \right| \right), \quad (5)$$

where $\Delta z = z_{\text{pred}} - z_{\text{true}}$, with z_{pred} and z_{true} indicating the predictions and true redshifts, respectively. Outlier percentage η demonstrates the fraction of severely inaccurate redshift predictions, and σ_{NMAD} is an accuracy metric that is robust against the outliers.

Figure 8 illustrates the results for the CNN. The accuracy σ_{NMAD} and outlier percentage η can reach 0.00047 and 0.954%, respectively. The accuracy successfully fulfills the requirement of $\sigma_{\text{NMAD}} < 0.005$ for BAO and other studies employing CSST slitless spectroscopic surveys (Y. Gong et al. 2019; H. Miao et al. 2024). The logarithmic SNR of *GI* bands are also displayed by the color bar. We notice that, as expected, the SNR decreases with respect to redshift and most outliers have relatively low SNRs. The redshift distribution is displayed in Figure 7, and is highly consistent with the true distribution. Furthermore, Figure 9 displays the accuracy and outlier percentage with respect to true redshifts in the upper and lower panel, respectively. The two metrics over the whole redshift range are also displayed in black dashed lines. As expected, both metrics remain steady at lower redshifts and becomes worse at higher redshifts.

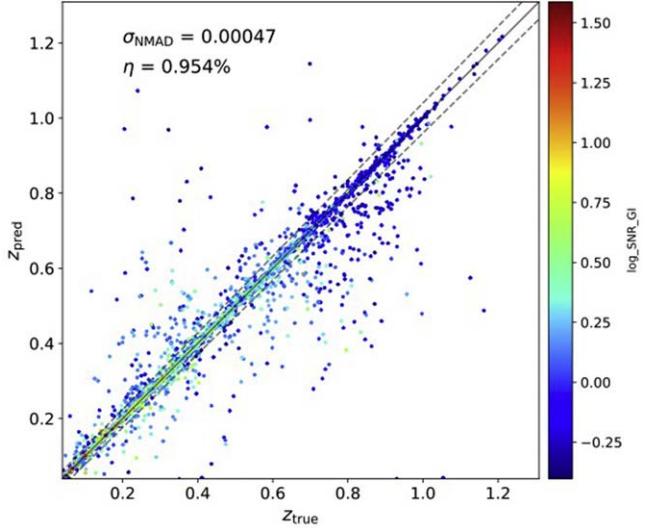


Figure 8. The results of the 1D CNN are illustrated, achieving accuracy $\sigma_{\text{NMAD}} = 0.00047$ and outlier percentage $\eta = 0.954\%$, respectively. The logarithmic SNRs of GI bands are indicated by the color bar.

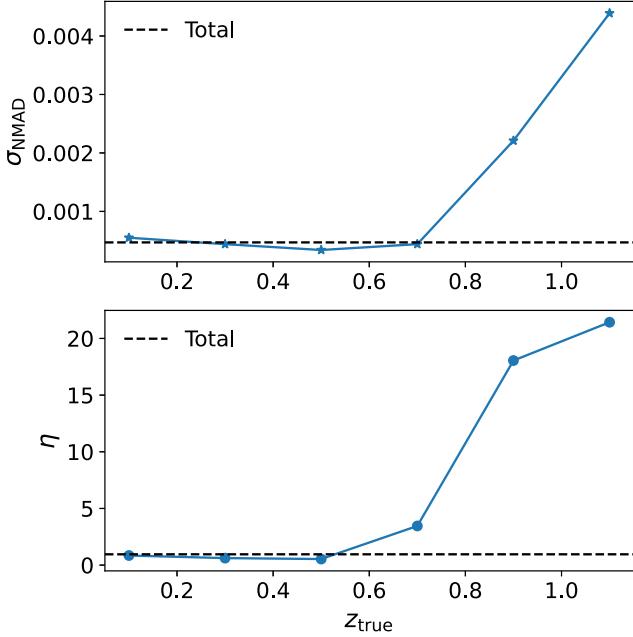


Figure 9. The accuracy σ_{NMAD} and outlier percentage η with respect to true redshifts are displayed in the upper and lower panels, respectively. The two metrics over the whole redshift range are also shown in black dashed lines.

As for the results from the BNN, apart from the two metrics mentioned above, we employ another metric to measure the performance of uncertainty predictions, i.e., the weighted mean uncertainty \bar{E} , which is defined as:

$$\bar{E} = \frac{\sum_i E_i / (1 + z_{i,\text{true}})}{N_{\text{total}}}, \quad (6)$$

where E_i is the uncertainty prediction for each source. The weight $1 + z_{i,\text{true}}$ applied to each source is to eliminate the bias from the evolution of redshifts. Figure 10 displays the reliability diagram for uncertainty predictions, from which we notice that the uncertainties are overestimated. After Beta calibration, the uncertainties better follow the statistical

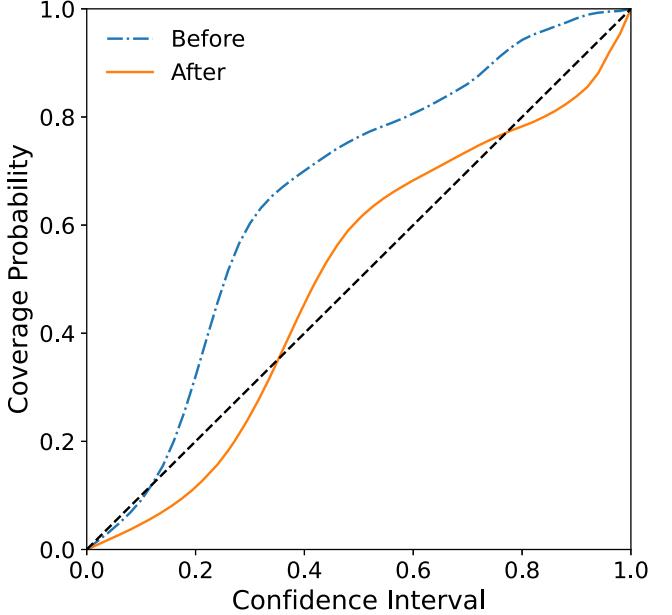


Figure 10. Reliability diagram for BNN results before and after the calibration. The black dashed line indicates that the uncertainties are well calibrated with the statistical principle perfectly followed.

principle as mentioned in Section 3.1. And Figure 11 shows the results after uncertainty calibration, where the error bars are displayed in light blue. Metrics σ_{NMAD} and η can achieve 0.00063 and 0.92%, respectively. Outlier percentage η slightly improves compared to point estimates illustrated in Figure 8, while the accuracy becomes a little worse, but still satisfies the requirement of cosmological studies. Furthermore, the weighed mean uncertainty \bar{E} can reach 0.00228. The redshift distribution is displayed in Figure 7, and similarly it is consistent with the true distribution. Figure 12 further analyzes the behavior of the uncertainties. The upper panel displays the weighted mean uncertainty \bar{E} with respect to the true redshift, in which the black dashed line shows the value over the whole redshift range. As expected, this metric similarly remains stable at lower redshifts and becomes worse as redshift increases. The lower panel shows the scatter plot between uncertainty and SNR in the GI band, and we notice that, with SNR increasing, the scatter of uncertainties decreases as expected.

For comparison, according to the data analysis pipeline for CSST slitless spectra, the traditional fitting for redshift estimations may produce an accuracy as low as $\sigma_{\text{NMAD}} \sim 0.01$ under such low SNR ~ 1 as shown in Figure 3. This demonstrates that a deep learning algorithm can significantly enhance the accuracy of redshift estimations, especially for low-SNR slitless spectra.

5. Discussion

As outlined in Section 2.1, the slitless spectra simulation software accepts SEDs and morphological parameters. The morphological parameters are employed to generate a Sérsic profile for each galaxy. This profile serves as the basis for dispersion. Nevertheless, not all galaxies have valid measurements of morphological parameters. Therefore, our selected sources are predominantly biased toward BGSS and LRGs at low redshifts, with relatively larger and brighter luminosities.

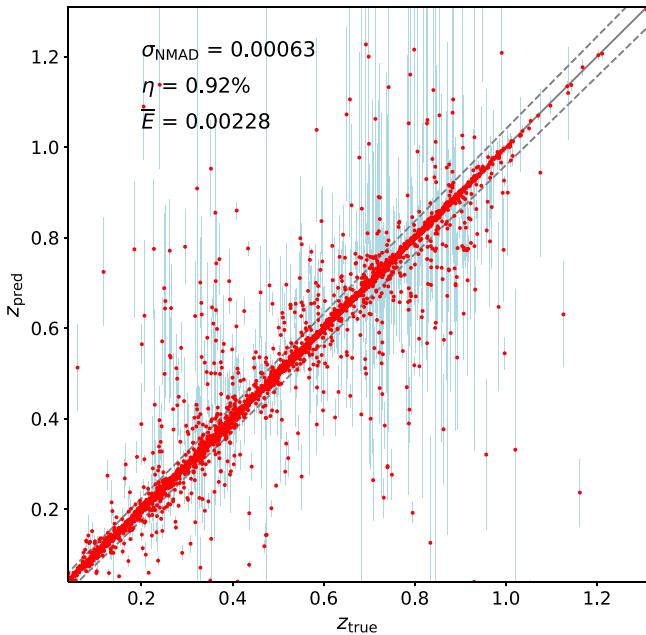


Figure 11. The results of the BNN after uncertainty calibration. The error bars are displayed in light blue. Over the whole redshift range, the BNN can reach $\sigma_{\text{NMAD}} = 0.00063$ and $\eta = 0.92\%$, respectively. And the weighted mean uncertainty \bar{E} can achieve 0.00228.

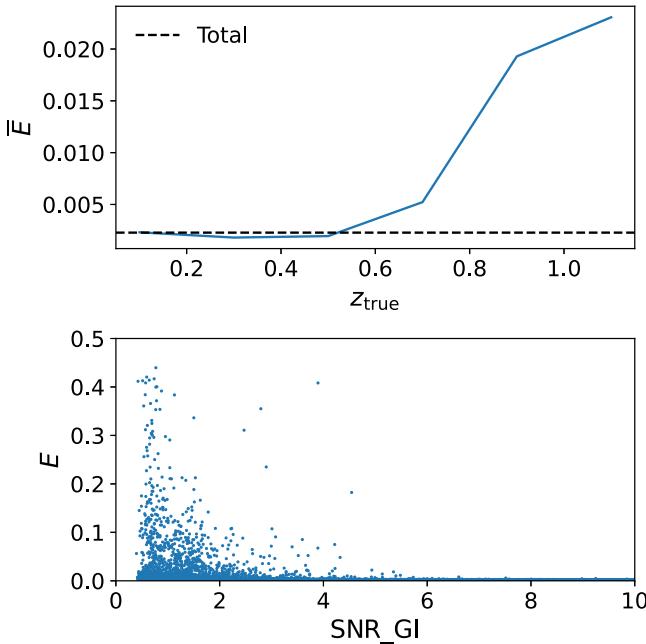


Figure 12. Upper: weighted mean uncertainty \bar{E} with respect to the true redshift. The value over the whole redshift range is also displayed in a black dashed line. Lower: weighted uncertainty E with respect to SNR in the GI band.

Emission line galaxies (ELGs), an ideal tracer at higher redshifts reaching ~ 1.6 , are unfortunately excluded due to inadequate morphological measurements. Consequently, the selected samples are unable to achieve the redshift limit of ~ 1.5 anticipated for the CSST spectroscopic survey, thereby hindering the comprehensive investigations of redshift accuracy by the deep learning method. Additionally, not all galaxies can be adequately represented by Sérsic profiles, particularly considering the significant uncertainties associated with some

morphological measurements. Furthermore, the morphologies of some galaxies may exhibit significant variations across different spectroscopic bands, rendering a single Sérsic profile inadequate. These biases and challenges can potentially be mitigated and addressed by incorporating a novel feature that directly simulates slitless spectra with galaxy images in reference photometric bands that are close to the slitless bands, such as F140W to G141 and F098W to G102 in the Hubble Space Telescope (M. Marinelli & L. Dressel 2024).

That said, the analysis of slitless spectra also presents challenges, particularly in terms of wavelength calibration and blending effects. These challenges severely impact the estimation of redshifts for individual galaxies. In the context of the CSST, wavelength calibrations are more complex compared to the Hubble Space Telescope (M. Marinelli & L. Dressel 2024) due to the absence of a direct image as a reference for determining wavelengths, considering the focal plane arrangement of the CSST. This can be addressed by employing CNNs to directly estimate redshifts from 2D spectral images. CNNs can analyze spectral images without wavelength information by leveraging knowledge gained from extensive training data. Further, blending effects are prevalent in spectroscopic observations, especially in deep surveys. Deblending algorithms, whether traditional or deep learning based, are actively being researched (C. J. Burke et al. 2019; B. Arcelin et al. 2021; S. Hemmati et al. 2022) and will be thoroughly investigated for slitless spectra in the future.

6. Conclusion

In this work, we employ a neural network to estimate the redshift from simulated slitless spectra in the CSST slitless spectroscopic survey. The simulation requires an SED and four morphological parameters including effective radius, Sérsic index, axis ratio, and position angle of each galaxy. To simulate the slitless spectra realistically, we use observational data from DESI-EDR and BOSS-DR16 with high-quality spectroscopic redshifts. The model spectra generated in the spectrum fitting process for these two observational data sets are considered as SEDs, and the sources are matched with DESI LS DR9 to retrieve the required morphological parameters. The SNRs of the slitless spectra are low with total SNRs peaking at ~ 1 , hence the key spectroscopic features used for redshift determinations are hard to identify. Therefore, we leverage the superior capability in processing noisy data of a neural network to estimate the redshifts from these slitless spectra.

Recognizing the importance of uncertainty predictions for several cosmological studies, we employ a Bayesian network to accomplish this task by providing redshift estimations along with uncertainties. To increase the robustness and converging speed, we construct the BNN based on a CNN for point estimates using transfer learning techniques. Gaussian random realizations are employed to largely augment the training size, ensuring the generative ability and noise tolerance of the BNN. After training, the uncertainty predictions for testing data are carefully calibrated. The BNN can achieve the results of $\sigma_{\text{NMAD}} = 0.00063$, $\eta = 0.92\%$, and $\bar{E} = 0.00228$, successfully satisfying the performance requirement of accuracy $\sigma_{\text{NMAD}} < 0.005$ for BAO and other cosmological studies based on slitless spectra of the CSST. Our approach can achieve a better performance than traditional SED fitting, particularly for low-SNR slitless spectra, serving as a complementary method for spectroscopic redshift estimation.

We also recognize that our work has certain limitations. These limitations can be attributed to our current version of the simulation software and the limited data size from the DESI early data release. With a newer version of the software and future data releases from DESI, we anticipate a comprehensive investigation of the spectroscopic redshift accuracy that can be achieved by deep learning algorithms for the CSST slitless spectroscopic survey.

Acknowledgments

X.C.Z. and Y.G. acknowledge the support from National Key R&D Program of China grant 2022YFF0503404, 2020SKA0110402, and the CAS Project for Young Scientists in Basic Research (No. YSBR- 092). This work is also supported by science research grants from the China Manned Space Project with grant Nos. CMS-CSST-2021-B01 and CMS-CSST-2021-A01. N.L. acknowledges the support from the science research grants from the China Manned Space Project (No. CMS-CSST-2021-A01), the CAS Project for Young Scientists in Basic Research (No. YSBR-062), and the Ministry of Science and Technology of China (No. 2020SKA0110100). H.Z. acknowledges the science research grants from the China Manned Space Project with Nos. CMS-CSST-2021-A02 and CMS-CSST-2021-A04 and the support from the National Natural Science Foundation of China (NSFC; grant Nos. 12120101003 and 12373010), National Key R&D Program of China (grant Nos. 2023YFA1607800, 2022YFA1602902), and Strategic Priority Research Program of the Chinese Academy of Science (grant Nos. XDB0550100).

ORCID iDs

- Yan Gong [ID](https://orcid.org/0000-0003-0709-0101) <https://orcid.org/0000-0003-0709-0101>
 Xuelei Chen [ID](https://orcid.org/0000-0001-6475-8863) <https://orcid.org/0000-0001-6475-8863>
 Run Wen [ID](https://orcid.org/0000-0002-8705-6327) <https://orcid.org/0000-0002-8705-6327>
 Yunkun Han [ID](https://orcid.org/0000-0002-2547-0434) <https://orcid.org/0000-0002-2547-0434>
 Hu Zou [ID](https://orcid.org/0000-0002-6684-3997) <https://orcid.org/0000-0002-6684-3997>
 Xian Zhong Zheng [ID](https://orcid.org/0000-0003-3728-9912) <https://orcid.org/0000-0003-3728-9912>
 Xiaohu Yang [ID](https://orcid.org/0000-0003-3997-4606) <https://orcid.org/0000-0003-3997-4606>
 Hong Guo [ID](https://orcid.org/0000-0003-4936-8247) <https://orcid.org/0000-0003-4936-8247>

References

- Agarap, A. F. 2018, arXiv:1803.08375
 Ahumada, R., Allende Prieto, C., Almeida, A., et al. 2020, *ApJS*, **249**, 3
 Arcelin, B., Doux, C., Aubourg, E., Roucelle, C. & LSST Dark Energy Science Collaboration 2021, *MNRAS*, **500**, 531
 Arendt, A. R., Perrott, Y. C., Contreras-Santos, A., et al. 2024, *MNRAS*, **530**, 20
 Bassett, B., & Hlozek, R. 2010, Dark Energy: Observational and Theoretical Approaches, arXiv:0910.5224
 Blundell, C., Cornebise, J., Kavukcuoglu, K., & Wierstra, D. 2015, arXiv:1505.05424
 Burke, C. J., Aleo, P. D., Chen, Y.-C., et al. 2019, *MNRAS*, **490**, 3952
 Busca, N., & Balland, C. 2018, arXiv:1808.09955
 Collister, A. A., & Lahav, O. 2004, *PASP*, **116**, 345
 Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, *AJ*, **145**, 10
 DESI Collaboration, Adame, A. G., Aguilar, J., et al. 2023, *AJ*, **168**, 58
 DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, arXiv:1611.00036
 Dey, A., Schlegel, D. J., Lang, D., et al. 2019, *AJ*, **157**, 168
 Dinh, L., Sohl-Dickstein, J., & Bengio, S. 2016, arXiv:1605.08803
 Euclid Collaboration, Mellier, Y., Abdurro'uf, et al. 2024, arXiv:2405.13491
 Fluri, J., Kacprzak, T., Lucchi, A., et al. 2022, *PhRvD*, **105**, 083518
 Gal, Y., & Ghahramani, Z. 2015, arXiv:1506.02142
 Gong, Y., Liu, X., Cao, Y., et al. 2019, *ApJ*, **883**, 203
 Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. 2017, arXiv:1706.04599
 Gupta, A., Zorrilla Matilla, J. M., Hsu, D., & Haiman, Z. 2018, *PhRvD*, **97**, 103515
 Hamilton, A. J. S. 1998, in The Evolving Universe, ed. D. Hamilton, 231 (Dordrecht: Kluwer Academic Publishers), 185
 He, K., Zhang, X., Ren, S., & Sun, J. 2015, arXiv:1512.03385
 He, Z., Er, X., Long, Q., et al. 2020, *MNRAS*, **497**, 556
 Hemmati, S., Huff, E., Nayyeri, H., et al. 2022, *ApJ*, **941**, 141
 Henghes, B., Thiyyagalingam, J., Pettitt, C., Hey, T., & Lahav, O. 2022, *MNRAS*, **512**, 1696
 Hortúa, H. J., García, L. Á., & Castañeda, C. L. 2023, *FrASS*, **10**, 1139120
 Hortúa, H. J., Volpi, R., Marinelli, D., & Malagò, L. 2020, *PhRvD*, **102**, 103509
 Hüllermeier, E., & Waegeman, W. 2021, *Mach Learn*, **110**, 457
 Ioffe, S., & Szegedy, C. 2015, arXiv:1502.03167
 Jimenez Rezende, D., & Mohamed, S. 2015, arXiv:1505.05770
 Jones, E., Do, T., Boscoe, B., et al. 2024, *ApJ*, **964**, 130
 Kingma, D. P., & Ba, J. 2014, arXiv:1412.6980
 Kiureghian, A. D., & Ditlevsen, O. 2009, *Structural Safety*, **31**, 105
 Kull, M., Filho, T. S., & Flach, P. 2017, in Proc. Machine Learning Research. 54, Proc. 20th Int. Conf. on Artificial Intelligence and Statistics, ed. A. Singh & J. Zhu (PMLR), 623, <https://proceedings.mlr.press/v54/kull17a.html>
 Le Fevre, O., Cassata, P., Cucciati, O., et al. 2013, *A&A*, **559**, A14
 Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. 1998, *Proceedings of the IEEE*, **86**, 2278
 Li, R., Napolitano, N. R., Tortora, C., et al. 2020, *ApJ*, **899**, 30
 Lilly, S. J., Le Fevre, O., Renzini, A., Zamorani, G., & Scudiego, M. 2007, *ApJS*, **172**, 70
 Louizos, C., & Welling, M. 2017, arXiv:1703.01961
 MacKay, D. J. C. 1995, *NCNS*, **6**, 469
 Mainieri, V., Anderson, R. I., Brinchmann, J., et al. 2024, arXiv:2403.05398
 Marinelli, M., & Dressel, L. 2024, WFC3 Instrument Handbook for Cycle 32, 16, 16
 Miao, H., Gong, Y., Chen, X., et al. 2024, *MNRAS*, **531**, 3991
 Min, Z., Xiao, X., Ding, J., et al. 2024, *PhRvD*, **110**, 063531
 Pan, S., Liu, M., Forero-Romero, J., et al. 2020, *SCPMA*, **63**, 110412
 Pasquet, J., Bertin, E., Treyer, M., Arnouts, S., & Fouchez, D. 2019, *A&A*, **621**, A26
 Pearson, W. J., Wang, L., Trayford, J. W., Petrillo, C. E., & van der Tak, F. F. S. 2019, *A&A*, **626**, A49
 Perreault Levasseur, L., Hezaveh, Y. D., & Wechsler, R. H. 2017, *ApJL*, **850**, L7
 Rastegarnia, F., Mirtorabi, M. T., Moradi, R., Vafaei Sadr, A., & Wang, Y. 2022, *MNRAS*, **511**, 4490
 Rezaei, S., McKean, J. P., Biehl, M., de Roo, W., & Lafontaine, A. 2022, *MNRAS*, **517**, 1156
 Sabelhaus, P. A., & Decker, J. E. 2004, *Proc. SPIE*, **5487**, 550
 Sadeh, I. 2014, in Statistical Challenges in 21st Century, 306, ed. A. Heavens, J.-L. Starck, & A. Krone-Martins (Cambridge: Cambridge Univ. Press), 316
 Schaefer, C., Geiger, M., Kuntzer, T., & Kneib, J. P. 2018, *A&A*, **611**, A2
 Schlegel, D., Dey, A., Herrera, D., et al. 2021, AAS Meeting Abstracts, **237**, 235.03
 Schlegel, D. J., Kollmeier, J. A., Aldering, G., et al. 2022, arXiv:2209.04322
 Schuldt, S., Suyu, S. H., Cafámeras, R., et al. 2021, *A&A*, **651**, A55
 Sérsic, J. L. 1963, *BAAA*, **6**, 41
 Sérsic, J. L. 1968, *Atlas de Galaxias Australes*
 Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. 2014, *Journal of Machine Learning Research*, **15**, 1929
 Tamura, N., Takato, N., Shimono, A., et al. 2016, *Proc. SPIE*, **9908**, 99081M
 Tewes, M., Kuntzer, T., Nakajima, R., et al. 2019, *A&A*, **621**, A36
 Treyer, M., Ait Ouahmed, R., Pasquet, J., et al. 2024, *MNRAS*, **527**, 651
 Wen, Y., Vicol, P., Ba, J., Tran, D., & Grosse, R. 2018, arXiv:1803.04386
 Zhan, H. 2018, in 42nd COSPAR Scientific Assembly, **42**, E1.16–4–18
 Zhang, Z., Shan, H., Li, N., et al. 2024, *A&A*, **683**, A209
 Zhou, X., Gong, Y., Meng, X.-M., et al. 2021, *ApJ*, **909**, 53
 Zhou, X., Gong, Y., Meng, X.-M., et al. 2022a, *MNRAS*, **512**, 4593
 Zhou, X., Gong, Y., Meng, X.-M., et al. 2022b, *RAA*, **22**, 115017
 Zou, H., Gao, J., Zhou, X., & Kong, X. 2019, *ApJS*, **242**, 8