China-VO Science Platform And Its Applications

Changhua Li Chenzhou Cui Yanxia Zhang Shirui Wei Maoyuan Huang

National Astronomical Observatories of China, CAS

Abstract: The Chinese Virtual Observatory (China-VO) has built a geographically distributed astronomical computing platform. This platform integrates dozens of astronomical data processing software tools that can be deployed on demand online, and features web-based interactive capabilities, enabling direct access to large-scale scientific data. The platform has already provided robust support and services for many astronomical research projects.

INTRODUCTION

The Science platform is a portal to build a private science environment for large scale data analysis and processing by a web browser. Fig. 1 shows the main interface, while Fig. 2 illustrates the integrated software tools.

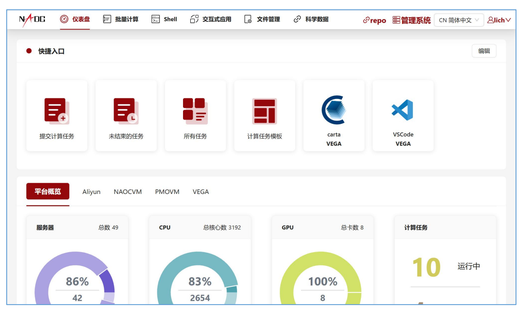

Fig. 1 The main interface

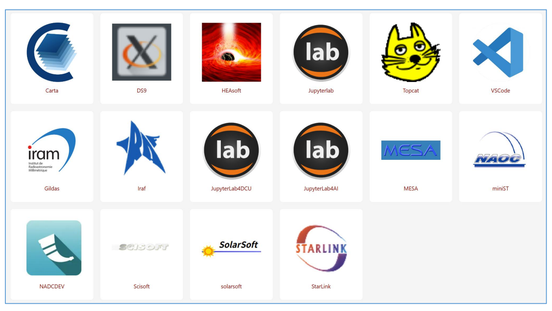

Fig. 2 List of the integrated software

The platform integrates a variety of commonly used datasets and corresponding software packages. Fig. 3 shows the access and processing of LAMOST data via pylamost.

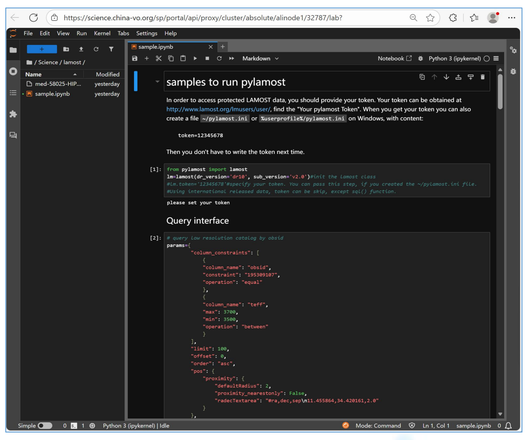

Fig. 3 The pylamost in jupyter

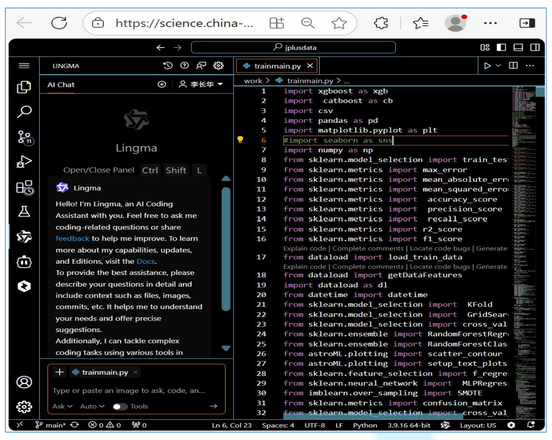

Fig. 4 AI assistant in the platform

The platform has built a unified authentication system. Users only need to register for an NADC Passport to access the platform's resources. The platform already supports a variety of scientific applications, with two examples listed below.

Case 1. A Photometric Redshift Catalogue of Galaxies from the DESI Legacy Imaging Surveys DR10 (Li et al. 2024)

Training Sample: DESI Legacy Imaging Surveys cover about 14000 deg2 in the Northern Galactic Cap and the South Galactic Cap region with g, r, z bands. We cross-matched data from multiple different spectroscopic redshift surveys with DESI Legacy Imaging Surveys DR10, obtaining 2,417,085 sources with spectroscopic redshifts.

By feature selection and optimization of model parameters, a optimal model was trained, the best performance (MSE = 0.0006, σ NMAD = 0.0098, and O = 0.08 per cent) is achieved, Fig. 5 shows the performance of model.

The galaxy catalog with photometric redshifts is released on Paperdata of NADC, with the access link: https://doi.org/10.12149/101485.

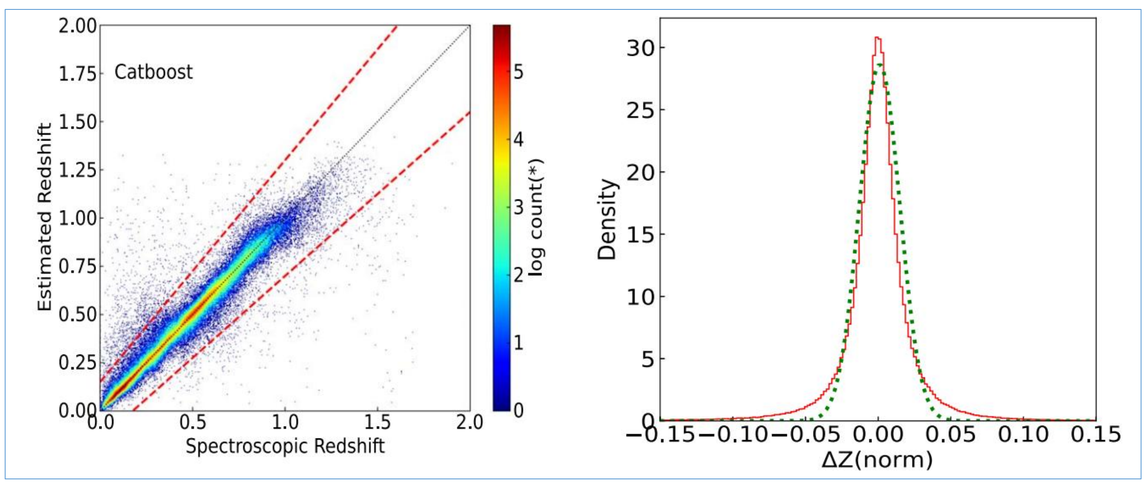

Fig. 5: The scatter figure of estimated photometric redshifts vs. spectroscopic redshifts and the Δz(norm) distribution by CatBoost.

The detailed information of model training can be refered Li et al. 2024, the doi is https://doi.org/10.3847/1538-3881/ad7c52.

Case 2. Photometric redshift estimation for emission line galaxies of DESI Legacy Imaging Surveys by CNN-MLP (Wei et al. 2025)

The Emission Line Galaxies (ELGs), which contain rich information about the processes occurring within galaxies (including star formation, ionisation/ionization, and chemical evolution), is a major focus of the DESI experiment, comprising one-third of the survey targets.

Emission-line galaxies with spectroscopic redshifts are cross-matched with DESI to generate training samples, with the total number of samples reaching 192, 375. Combining catalog data with image data, each source includes image data from 10 channels and 85 optimal photometric features. Fig. 6 shows the network architecture of the CNN-MLP model.

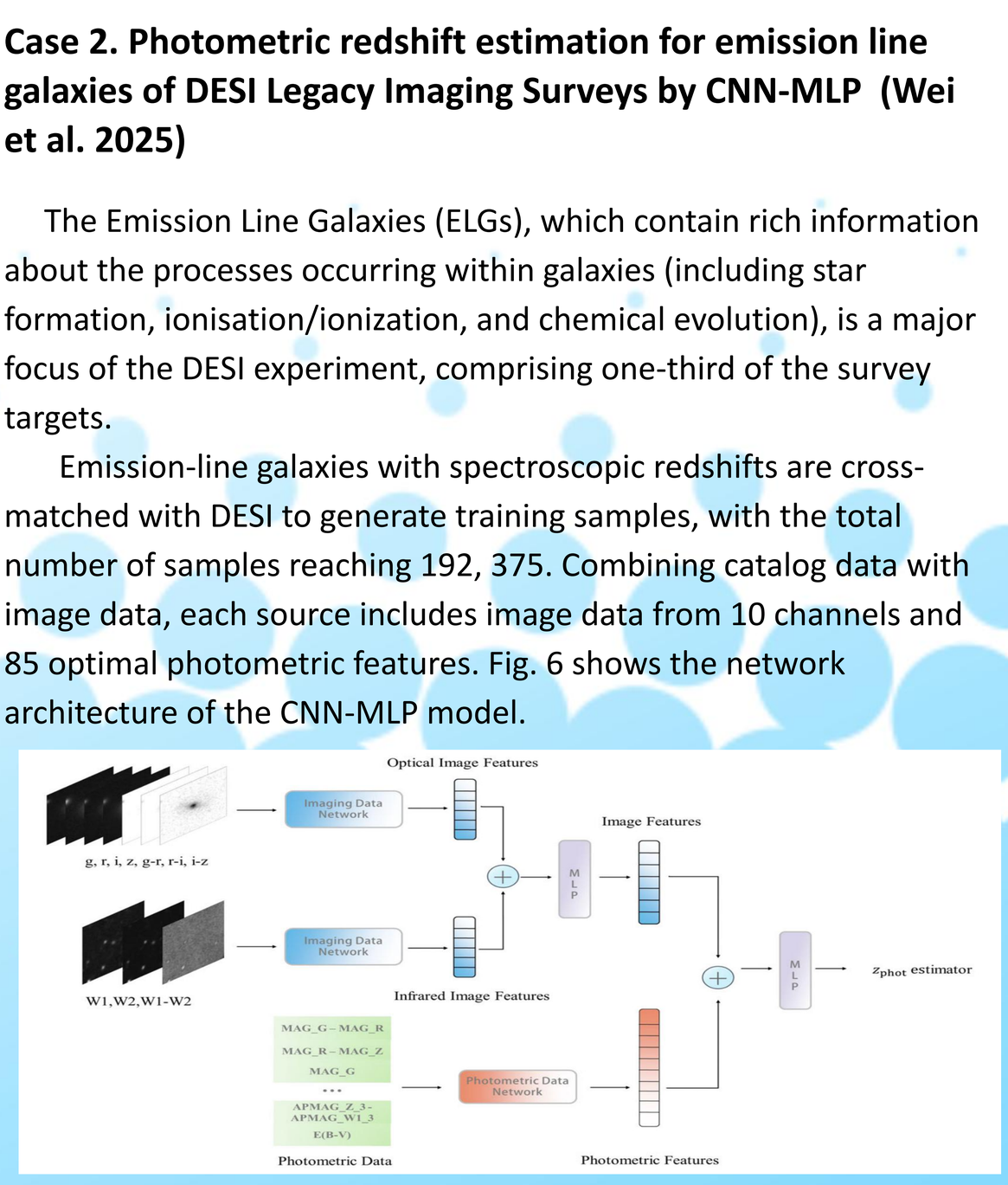

Fig. 6: Schematic diagram of the CNN-MLP model

The performance of CNN-MLP model reach σ NMAD = 0.0098, and O = 0.08 per cent. The detailed information of model can be refered Wei et al. 2025, the paper doi is https://doi.org/10.1017/pasa.2025.10056.

https://nadc.china-vo.org Contact me at: lich@nao.cas.cn