Postprint of Gas-Phase Metallicity Estimation for LAMOST Star-Forming Galaxies Based on Deep Learning

Authors: Feng Jiabao, Wang Lili, Li Ping

Date: 2026-04-10T16:45:03+00:00

Abstract

Gas-phase metallicity is a crucial physical quantity for measuring the chemical evolution of star-forming galaxies, and its estimation is essential for a deep understanding of galaxy formation and evolution processes. Traditional methods for measuring gas-phase metallicity rely on the intensity ratios of specific emission lines, which involve complex data processing and are difficult to adapt to the automation requirements of large-scale sky survey data. To address this, a deep learning model based on a Convolutional Neural Network (CNN) is proposed. Using the full spectra observed by LAMOST (Large Sky Area Multi-Object Fiber Spectroscopic Telescope) as input, the model achieves automated estimation of gas-phase metallicity for star-forming galaxies without the need for redshift correction or spectral line measurement. The model consists of eight 1D convolutional layers, four max-pooling layers, and one fully connected layer, learning the nonlinear mapping between spectra and metallicity through regression. Experimental results show that the model is fundamentally consistent with traditional methods, with an error of 0.0829 dex. Furthermore, the CNN model demonstrates good robustness across different signal-to-noise ratios and redshift ranges, and the Mass-Metallicity Relation (MZR) established using the metallicity predicted by the model is largely consistent with the empirical MZR. Finally, the trained model was applied to the LAMOST Data Release 10 (DR10) Low-Resolution Survey (LRS), constructing a gas-phase metallicity catalog containing approximately 20,000 star-forming galaxies. This catalog is accessible via the Science Data Bank of the Chinese Academy of Sciences (https://www.scidb.cn/s/UVBRzm).

ChinaRxiv 中X

Full Text

Preamble

Vol. 67 No. 2 March 2026

ACTA ASTRONOMICA SINICA

67 No. 2

Mar., 2026

doi: 10.15940/j.cnki.0001-5245.2026.02.010

Gas-Phase Metallicity Estimation of LAMOST Star-Forming Galaxies Based on Deep Learning*

Jiabao Feng[1,2], Lili Wang[2†], Ping Li[1,2] (1 School of Information and Control Engineering, Jilin Institute of Chemical Technology, Jilin 132022) (2 School of Computer and Information, Dezhou University, Dezhou 253023)

摘要

Gas-phase metallicity is a critical physical quantity for measuring the chemical evolution of star-forming galaxies. Accurately estimating this parameter is essential for gaining a profound understanding of the processes governing galaxy formation and evolution. Traditional methods for measuring gas-phase metallicity rely on the intensity ratios of specific emission lines; however, these approaches involve complex data processing and struggle to adapt to the demands of large-scale sky surveys.

To address the growing demand for automation in processing astronomical data, this paper proposes a deep learning approach based on Convolutional Neural Networks (CNNs).

The model utilizes the full spectra observed by LAMOST (Large Sky Area Multi-Object Fiber Spectroscopic Telescope) as input. In the absence of additional information, it leverages deep learning techniques to extract complex spectral features. By training on high-quality labeled data, the model can effectively map these features to stellar atmospheric parameters, such as effective temperature (T_{eff}), surface gravity ($\log g$), and metallicity ([Fe/H]).

[Figure 1: see original paper]

The architecture of the proposed neural network is designed to handle the high dimensionality and noise characteristics inherent in LAMOST spectra. Specifically, the model employs a series of convolutional layers to capture localized spectral lines and broad continuum features simultaneously. This approach ensures that both the subtle absorption features and the overall spectral energy distribution are accounted for during the parameter estimation process.

To evaluate the performance of the model, we conducted a series of validation tests using a subset of the LAMOST data with known parameters from high-resolution spectroscopic surveys. The results indicate that the deep learning model achieves a high degree of accuracy and precision, significantly reducing the systematic errors often associated with traditional template-matching methods. Furthermore, the model demonstrates robust performance even for spectra with relatively low signal-to-noise ratios (SNR), making it a valuable tool for large-scale galactic surveys.

Under the premise of requiring redshift and spectral line measurements, this study achieves the automated estimation of gas-phase metallicity in star-forming galaxies. The model consists of eight 1-D convolutional layers and four···

The architecture consists of a max-pooling layer and a single fully connected layer, which learns the non-linear mapping between the spectra and metal abundances through a regression approach. Experimental results demonstrate that ···

The results are fundamentally consistent with traditional methods, yielding an error of 0.0829 dex. Furthermore, the CNN model demonstrates high robustness across a wide range of signal-to-noise ratios and redshifts.

ity, and the Mass-Metallicity Relation (MZR) established using the model-predicted metallicities is fundamentally consistent with the empirical MZR.

[Figure 1: see original paper]

The chemical evolution of galaxies is a core component of galaxy formation and evolution studies. The MZR, which describes the correlation between the stellar mass of a galaxy and its gas-phase or stellar metallicity, serves as a critical constraint for theoretical models. In this work, we utilize machine learning techniques to predict the metallicities of a large sample of galaxies. By comparing our results with established empirical relations, we demonstrate that the deep learning model not only captures the global trends of chemical enrichment but also maintains high precision across different mass regimes.

As shown in [Figure 1: see original paper], the distribution of predicted metallicities follows the expected physical scaling laws. Furthermore, the statistical properties summarized in indicate that the residuals between our model predictions and the observational benchmarks are minimized. This consistency validates the robustness of our approach in reconstructing the chemical enrichment history of galaxies based on their integrated physical properties.

consistent. Finally, the trained model was applied to the low-resolution survey data from the LAMOST Tenth Data Release (DR10).

Abstract

We have constructed a gas-phase metallicity catalog comprising approximately 20,000 star-forming galaxies using data from the LAMOST (Large Sky Area

Multi-Object Fiber Spectroscopic Telescope) Low-Resolution Survey (LRS). This catalog is accessible through the Chinese Virtual Observatory (China-VO) and provides a significant resource for studying the chemical evolution of galaxies.

Introduction

The gas-phase metallicity of galaxies serves as a crucial indicator of their evolutionary history, reflecting the complex interplay between star formation, gas accretion, and galactic outflows. By analyzing the chemical enrichment patterns across large samples of galaxies, we can better understand the fundamental processes governing galaxy formation and evolution. The LAMOST survey, with its vast spectroscopic database, offers an unprecedented opportunity to conduct such large-scale statistical studies.

Data and Methodology

The primary data source for this study is the LAMOST Low-Resolution Survey. We selected star-forming galaxies based on their spectral characteristics, specifically focusing on those with high signal-to-noise ratios in key emission lines. To determine the gas-phase metallicity, we employed standard diagnostic methods, including the use of strong-line ratios.

Sample Selection

Our initial sample was drawn from the LAMOST DR series, filtering for objects classified as galaxies. We applied several quality control criteria: 1. A signal-to-noise ratio (SNR) in the r-band greater than 10. 2. Clear detection of essential emission lines such as $H\alpha$, $H\beta$, $[OIII]\lambda5007$, and $[NII]\lambda6584$. 3. Classification as star-forming galaxies using the BPT (Baldwin-Phillips-Terlevich) diagram to exclude Active Galactic Nuclei (AGN) contamination.

[Figure 1: see original paper]

Metallicity Estimation

The metallicity, expressed as $12+\log(O/H)$, was calculated using well-calibrated empirical and theoretical relations. We utilized the O3N2 and N2 indices as defined by [**?**] to ensure consistency with previous literature. The resulting catalog provides not only the metallicity values but also the associated uncertainties and the equivalent widths of the primary emission lines.

Results and Discussion

The final catalog contains approximately 20,000 star-forming galaxies, spanning a wide range of stellar masses and redshifts. Our analysis

The scientific data associated with this study can be accessed via the Science Data Bank of the Chinese Academy of Sciences at the following URL: https://www.scidb.cn/s/UVBRzm.

关键词

Galaxies: gas-phase metallicity, Methods: data analysis, Methods: statistical, Methods: deep learning

CLC number: P157; Document code: A

1 引言

Star-forming galaxies are among the most active types of galaxies in the universe. Their formation and evolutionary processes are closely linked to star formation activity and the physical properties of their gas [?]. As a key parameter for measuring the content of heavy elements within the interstellar medium, gas-phase metallicity not only influences physical mechanisms such as the star formation rate and gas cooling processes but also reflects the chemical evolution history of the galaxy [?]. Consequently, the precise determination of gas-phase metallicity in star-forming galaxies is of great importance for understanding galactic chemical evolution.

Gas-phase metallicity is typically expressed in terms of oxygen abundance. Higher metallicity generally indicates that a galaxy has undergone multiple cycles of star formation and supernova explosions, thereby enriching the medium with more heavy elements [?]. Traditionally, estimation methods for gas-phase metallicity primarily include the direct method (also known as the electron temperature method, denoted as the T_e-method) [?] and the strong-line method [?]. The direct method calculates the electron temperature by measuring the intensity ratios of specific emission lines (such as [OIII] λ4363, [NII] λ5755, and [SII] λ6312, where brackets denote forbidden lines and λ represents wavelength) to derive the gas metallicity. In contrast, the strong-line method estimates metallicity based on empirical relationships established using strong emission lines (for example, R_{23} [?], $O3N2$ [?], etc.).

Received original manuscript April 3, 2025; received revised manuscript June 4, 2025. † jsjxwll@126.com

Te(cid:21)(cid:21)(cid:21)(cid:21)

ACTA ASTRONOMICA SINICA

Various methods for estimating gas-phase metallicity have distinct ranges of applicability, such as $N2O2$ [?] and $N2$ [?]. The R_{23} method proposed by Pagel et al. [?] is suitable for low to intermediate metallicity estimations, while the $O3N2$ and $N2$ methods proposed by Alloin et al. [?] are more appropriate for intermediate metallicity measurements. In contrast, the $N2O2$ method proposed by Kewley et al. [?] exhibits greater stability at high metallicities. Furthermore, Tremonti et al. [?] employed a Bayesian spectral fitting approach to analyze 53,000 galaxy spectra from the Sloan Digital Sky Survey (SDSS). Utilizing the models developed by Charlot and Longhetti in 2001 [?]—which combine stellar population synthesis with photoionization calculations—they estimated gas-phase metallicities by fitting multiple strong emission lines and established the mass-metallicity relation.

However, with the rapid growth of astronomical observational data, traditional methods face challenges regarding high computational complexity when processing large-scale datasets. In this context, machine learning and deep learning technologies have been widely applied in astronomical research due to their advantages in efficiently processing massive data and automatically extracting complex features. Hoyle et al. [?] trained machine learning models to estimate redshifts for SDSS galaxy spectra, improving both accuracy and robustness. Wang et al. [?] employed Convolutional Neural Networks (CNNs) to train on low-redshift SDSS galaxy spectra to predict stellar age, metallicity, extinction, and velocity dispersion. Regarding gas-phase metallicity estimation, Ho et al. [?] trained a model based on a Multilayer Perceptron (MLP) that utilizes emission line intensity ratios from 950 HII regions to accurately predict oxygen abundance and reconstruct metallicity gradients. Teimoorinia et al. [?] proposed a method based on a Random Forest regression model to re-evaluate the transformation relationships between different strong-line metallicity calibrations using strong emission line features from SDSS spectroscopic data. These methods rely on the individual measurement of specific emission line intensities and establish mappings between metallicity and emission lines through empirical relations or machine learning models.

These studies demonstrate that machine learning and deep learning technologies have shown powerful capabilities across multiple astronomical tasks, including the prediction of redshift, metallicity, and stellar population parameters.

The Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST) is a wide-field, multi-object spectroscopic survey facility independently developed by China. It possesses the capability for high-efficiency, large-scale spectroscopic acquisition and has accumulated a massive volume of spectroscopic data since it began operations in 2011 [?]. Its tenth data release (Data Release 10, DR10) was officially opened to the international astronomical community in 2024.

The release includes two types of spectroscopic data: the Low-Resolution Survey (LRS) and the Medium-Resolution Survey (MRS). The LRS data contains 11,441,011 spectra, including 261,381 galaxy spectra, covering a wavelength range of 3700–9000 Å with a resolution of approximately 1800. This study proposes a method based on Convolutional Neural Networks to calculate the gas-phase metallicity of star-forming galaxies in LAMOST DR10 LRS. Compared to traditional methods, this approach uses the full spectrum directly as input, requires no redshift correction, and avoids the individual calculation of specific emission line intensities, thereby reducing uncertainties introduced during the data preprocessing stage.

2 数据

This study utilizes 261,381 galaxy spectra observed by LAMOST DR10, which were filtered according to the following criteria. To ensure that key emission lines (such as H_α, $[OIII]$, $[NII]$, and H_β) remain within the observable range, the redshift range was restricted to $z < 0.25$. Additionally, to ensure data quality, a minimum r-band signal-to-noise ratio (S/N_r) of 3 was required. Following these criteria, 242,500 galaxy spectra were selected.

The Max Planck Institute for Astrophysics-Johns Hopkins University (MPA-JHU) provides key parameters such as gas-phase metallicity based on SDSS data. We selected the median gas-phase metallicity (oh_p50) [**?**] from this catalog as the target variable for our regression model. To obtain labels for the regression model, we performed a cross-match between the LAMOST and SDSS MPA-JHU DR8 catalogs based on right ascension (ra) and declination (dec), using a matching radius of 1'' . Among the 242,500 LAMOST galaxy spectra, approximately 151,000 (62.28%) were successfully matched to the MPA-JHU catalog, while the remaining 91,400 (37.72%) lacked corresponding parameter information. From the successfully matched data, we further selected star-forming galaxies (bpt-class=1) and excluded outliers (oh_p50 = -9999), ultimately obtaining approximately 38,500 LAMOST star-forming galaxy spectra with available metallicity measurements. [Figure 1: see original paper] shows the distribution of redshift, r-band signal-to-noise ratio (S/N_r), and r-band Petrosian magnitude (petroMag_r) for the selected data. It can be observed that the redshifts are primarily concentrated between [0.02, 0.15], S/N_r is mainly distributed between [3, 30], and petroMag_r falls predominantly within the range of [16, 18].

Prior to model training, the selected data underwent preprocessing. First,

(cid:12)(cid:11)0:002 z 0:3(cid:0)

Feng Jiabao et al.: Estimating Gas-Phase Metallicity of LAMOST Star-Forming Galaxies Based on Deep Learning

All spectra were unified to a wavelength range of 3800 Å to 8000 Å using linear interpolation, with an interpolation step size set to 1 Å. This process ensures that each spectrum consists of exactly 4201 data points. Subsequently, a min-max normalization

method was applied to scale the spectral flux into the [0, 1] interval. This normalization ensures comparability between the model input data and provides a reliable data foundation for subsequent model training.

1 Distributions of redshift, signal-to-noise ratio in the r-band (S/N_r), and Petrosian magnitude in the r-band

(petroMag_r) in the dataset

3 方法

Formal Representation.

This chapter introduces the design and training of a Convolutional Neural Network (CNN) model for predicting gas-phase metallic abundance. The model is designed to automatically extract deep features from spectroscopic data and map these features to the target values through regression.

3.1 网络架构

The architecture of the CNN model constructed in this study is illustrated in Figure 2 [Figure 2: see original paper]. It consists of an input layer, multiple convolutional and pooling layers, and a final output layer. Through a hierarchical feature extraction mechanism, the entire network achieves a non-linear mapping from the raw spectra to metal abundances. The specific details are described as follows:

1．输入层：模型的输入为预处理后的光谱数

The dataset consists of spectral data where each individual spectrum contains 4,201 data points. The corresponding label data for these spectra is the gas-phase metallicity, denoted as oh_p50.

2．网络层：由于光谱数据为一维数据，我们采

One-dimensional convolutional layers (Conv1D) and one-dimensional max-pooling layers (MaxPooling1D) are employed. The CNN model consists of eight convolutional layers, four max-pooling layers, one flatten layer, and one fully connected layer.

Let Conv1Dxx denote a layer where xx represents the number of channels. Using the pooling layers as boundaries, the CNN model can be divided into four modules. Within each module, the number of channels in the convolutional layers remains constant, sequentially increasing to 16, 32, 64, and 128 across modules. This architecture facilitates the extraction of increasingly complex features as the network depth increases.

The kernel size for all convolutional layers is set to 3 with a stride of 1. For the max-pooling layers, the pooling window size is 3 with a stride of 3. The output of each layer is···

2 Architecture of the convolutional neural network

108642000.050.100.150.200.250.30Density0.080.060.040.020Density0.60.40.20Densityredshift3102030405015161
Pool1D161399*1Conv1D321397*1Conv1D321395*1Conv1D644463*1Conv1D128151*1Conv1D644461*1Conv1D128149
*Pool1D128449*1Max-Pool1D641553*1Max-Pool1D324465*1Flatten6272*1Fully Connected1*1Output

ChinaRxiv 中国X

ACTA ASTRONOMICA SINICA

1 Introduction

In recent years, the rapid development of machine learning and deep learning has significantly impacted the field of astronomy. With the advent of large-scale sky surveys, the volume of astronomical data has grown exponentially, necessitating more efficient and automated methods for data processing and analysis. Traditional manual classification and feature extraction methods are no longer sufficient to meet the demands of modern astronomical research.

2 Methodology

The core of our approach involves the application of advanced neural network architectures to classify celestial objects and identify transient events. By utilizing convolutional neural networks (CNNs) and recurrent neural networks (RNNs), we can extract complex spatial and temporal features from multi-wavelength imaging and light curve data.

2.1 Data Preprocessing Before feeding the data into the models, several preprocessing steps are required. These include noise reduction, normalization, and data augmentation to ensure the robustness of the training process. For spectroscopic data, we apply redshift corrections and continuum subtraction to isolate the relevant emission and absorption features.

2.2 Model Architecture We employ a deep residual network (ResNet) as the backbone for image classification. The architecture is defined by the following mapping:

$$y = \mathcal{F}(x, \{W_i\}) + x$$

where x and y are the input and output vectors of the layers considered, and the function $\mathcal{F}(x, \{W_i\})$ represents the residual mapping to be learned. This structure allows for the training of much deeper networks by addressing the vanishing gradient problem.

[Figure 1: see original paper]

3 Results and Discussion

Our experimental results demonstrate that the proposed deep learning model achieves high accuracy in the classification of various types of galaxies and stars. As shown in , the precision and recall rates exceed those of traditional support vector machine (SVM) and random forest classifiers.

Furthermore, the model' s performance on the test dataset indicates strong generalization capabilities. We observed that the integration of multi-band photometric data significantly improves the classification of high-redshift quasars. The relationship between the predicted and observed magnitudes is given by:

$$m_{pred} = \alpha \cdot m_{obs} + \beta$$

where α and β are coefficients determined during the calibration phase.

4 Conclusion

This study highlights the potential of machine learning in revolutionizing astronomical data analysis

3．激活函数：所有卷积层均使用修正线性单

The Rectified Linear Unit (ReLU) is employed as the activation function to introduce non-linearity, thereby enhancing the model' s ability to fit complex relationships.

4．输出层：经过展平操作后，数据输入到全连

connection layer, which ultimately outputs the predicted gas-phase metallicity.

5. Loss Function: Mean Squared Error (MSE) is employed as the loss function to calculate the average squared difference between the predicted values and the ground truth. The objective during the training process is to minimize this loss function to improve prediction accuracy. The formula for calculating MSE is as follows:

and the ground truth (

prediction performance. The specific evaluation includes the following aspects: Section 4.1 analyzes the discrepancies within the CNN model; Section 4.2 compares the prediction performance of the CNN against commonly used models, such as Random Forest and Multi-Layer Perceptron (MLP); Sections 4.3 and 4.4 analyze the impact of the signal-to-noise ratio (SNR) and redshift on model performance, respectively; and Section 4.5 explores the model's ability to reproduce the mass-metallicity relationship.

where n represents the number of samples;

is the ground truth for the i-th sample.

is the predicted value for the i-th sample.

6. Optimizer and Learning Rate Scheduling: The Adam optimizer is utilized, as it combines an adaptive learning rate with a momentum mechanism to effectively enhance the convergence speed and stability of the model. The initial learning rate is set to 0.001, and a `ReduceLROnPlateau` scheduler is implemented. If the validation loss does not improve for 10 consecutive

training epochs, the learning rate is automatically reduced by a factor of 10. This approach improves the stability of the model and enhances its ability to search for the global optimum.

3.2 网络训练

To train and evaluate the model, we divided approximately 38,500 samples into a training set, a validation set, and a test set according to a ratio of 7:1:2. The training set is used to optimize the model parameters, ensuring that the predicted values are as close as possible to the ground truth. The validation set is employed for real-time performance evaluation and hyperparameter tuning, while the test set is reserved for the final assessment of the model's generalization capability. [Figure 3: see original paper] illustrates the distribution of true metallicities across the training, validation, and test sets. The distributions are highly consistent across all three sets, providing a reliable foundation for model training and evaluation.

During the training process, each batch consists of 128 spectral data points and their corresponding labels. The model utilizes the backpropagation algorithm to update its parameters, aiming to minimize the loss function.

Training is conducted for a total of 100 epochs. The optimal model is selected by monitoring the loss on the validation set and is subsequently saved for final testing.

4 模型评估

Evaluation of Model Performance

This chapter evaluates the predictive performance of the constructed Convolutional Neural Network (CNN) model on the test dataset. To ensure a comprehensive assessment of the model' s effectiveness, we utilize several standard metrics common in machine learning research, including accuracy, precision, recall, and the F1-score. These metrics provide a multi-dimensional view of the model' s ability to generalize to unseen data.

[Figure 1: see original paper]

As illustrated in [Figure 1: see original paper], the training process demonstrates a steady convergence of the loss function, suggesting that the model successfully captures the underlying patterns within the data without significant overfitting. The evaluation on the test set serves as a critical benchmark for determining the practical utility of the proposed architecture in real-world scenarios.

Performance Metrics Analysis

The quantitative results obtained from the test set are summarized in . The model achieves an overall accuracy of $\mathcal{A}$, which indicates a high degree of reliability in its classification capabilities. Furthermore, the precision and recall values suggest that the CNN maintains a robust balance between minimizing false positives and ensuring that relevant instances are correctly identified.

To further analyze the model' s behavior, we examine the relationship between the input features and the resulting predictions. By applying the transformation $\mathcal{F}(x) = y$, where x represents the input vector and y the predicted class, we can observe how the convolutional layers extract hierarchical features. The mathematical formulation of the final output layer is given by:

$$\hat{y} = \sigma(W \cdot h + b)$$

where W denotes the weight matrix, b is the bias vector, and σ represents the activation function. This evaluation confirms that the architectural choices, such as the selection of specific kernel sizes and pooling strategies, contribute significantly to the overall performance.

3 Distribution of true values of gas-phase metallicity

for training, validation, and test sets

4.1 CNN 模型预测结果分析

This section analyzes the discrepancy between the gas-phase metallicities predicted by the CNN model and their corresponding ground-truth values. [Figure 4: see original paper] provides a visual comparison between the CNN model' s predicted values and the true values. In this figure, the red median line closely aligns with the black 1:1 ideal line. Furthermore, the mean and standard deviation of 0.0829 on the test set demonstrate that the model achieves high overall predictive accuracy.

It is worth noting that the prediction errors are relatively larger in the low-metallicity ($12 + \log(\mathrm{O/H}) < 8.2$) and high-metallicity ($12 + \log(\mathrm{O/H}) > 9.0$) regimes. This is likely due to the limited number of samples available in these specific ranges (as shown in [Figure 3: see original paper]). Overall, the CNN model demonstrates high accuracy and stability in the regression task of estimating gas-phase metallicity.

4.2 不同模型性能对比

This section compares the predictive performance of the CNN model with two classic models: Random Forest (RF) regression and the Multi-Layer Perceptron (MLP).

RF is an ensemble learning method that improves robustness and reduces overfitting by constructing multiple decision trees and averaging their predictive results. Each tree is trained on a random subset of the data, and the final prediction is obtained by averaging the outputs of all individual trees. In this study, the RF

1https://pytorch.org/docs/stable/generated/torch.optim.lr_{scheduler}.ReduceLROnPlateau.html

ZpredZtrueMSE=1NN∑i=1(Zpred;i(cid:0)Ztrue;i)2;(1)NZpred;iiZtrue;iiZpredZtrueTrain SetValidation SetTest SetDensity2.52.01.51.00.507.88.08.28.48.68.89.09.29.4Gas-phase Metallicity (oh_p50)((cid:22))(cid:0)0:0002(cid:27)<8:5>9:2

The model consists of 100 decision trees ($ntree = 100$), achieving an optimal balance between computational efficiency and predictive performance. The Multi-Layer Perceptron (MLP) is a feedforward artificial neural network comprising an input layer, multiple hidden layers, and an output layer. Each neuron is fully connected to all neurons in the subsequent layer, enabling the network to learn complex non-linear relationships. In this study, the MLP architecture includes an input layer, an output layer, and three hidden layers, with each hidden layer containing 1024 neurons to extract spectral features.

We trained the CNN, RF, and MLP models on the same dataset and evaluated their performance on the test set. summarizes the key metrics for the three models. It can be observed that the CNN model yields the smallest error, indicating that the CNN possesses a significant advantage over the other two models.

) and standard deviation (

CNN, RF, and MLP models

Model

Mean ( )

Standard Deviation ( )

−0.0002

−0.0053

4.3 不同信噪比下的模型评估

The signal-to-noise ratio (SNR) may influence the prediction accuracy of the CNN model. To analyze the model' s performance under varying noise levels, we partitioned the data into several intervals based on the r-band signal-to-noise ratio (S/N_r): $[3, 5)$, $[5, 7)$, $[7, 10)$, $[10, 15)$, $[15, 25)$, and $[25, +\infty)$. We then calculated the standard deviation for each interval (see [Figure 5: see original paper] (a)).

The results indicate that the prediction error is highest when S/N_r is below 5, with the standard deviation approaching 0.11 dex. As S/N_r increases, the prediction error gradually decreases, falling to approximately 0.05 dex at higher signal-to-noise levels. These findings demonstrate that the prediction accuracy of the CNN model improves accordingly as the quality of the input data increases.

4 Comparation of the predicted gas-phase metallicity

values using our CNN model with the true values

5 Model performance across different intervals: signal-to-noise ratio (a) and redshift (b)

1:1 LineMedian LineUpper 1σ LineLower 1σ Lineμ: -0.0002σ: 0.082917.515.012.510.07.55.02.57.88.08.28.48.68.
density (a.u.)(cid:22)=(cid:0)0:0002(cid:27)=0:0829(cid:22)(cid:27)Zpred(cid:0)Ztrue 250.120.100.080.060.0451
(Zpred−Ztrue)0.120.100.080.060.04σ (Zpred−Ztrue)S/N_r00.050.100.150.200.25Redshift(a)(b)

Acta Astronomica Sinica

Abstract

In recent years, the rapid development of machine learning and deep learning has provided powerful tools for processing and analyzing massive amounts of astronomical data. This paper reviews the current status and progress of applying these technologies to various fields of astronomy, including galaxy morphology classification, stellar spectral analysis, exoplanet detection, and transient source identification. We discuss the advantages of different algorithmic architectures—such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Generative Adversarial Networks (GANs)—in handling multi-dimensional and multi-wavelength observational data. Furthermore, we address the challenges currently faced by these methods, such as the interpretability of "black box" models, the requirement for high-quality labeled training sets, and the generalization capabilities of models across different observational facilities. Finally, we provide an outlook on the future directions of machine learning in the era of Large Synoptic Survey Telescopes (LSST) and other next-generation survey projects.

1 Introduction

Astronomy has entered the era of big data. With the operation of large-scale survey projects such as the Sloan Digital Sky Survey (SDSS), the Gaia mission, and the upcoming Vera C. Rubin Observatory, the volume, velocity, and complexity of astronomical data have increased exponentially. Traditional manual analysis and classical statistical methods are no longer sufficient to meet the demands of modern astronomical research. Consequently, the integration of artificial intelligence, particularly machine learning and deep learning, has become an inevitable trend in the development of contemporary astronomy.

Machine learning algorithms are capable of automatically extracting features from high-dimensional data, identifying complex patterns, and performing efficient classification and regression tasks. In the field of image processing, deep learning models have demonstrated performance surpassing that of human experts in tasks such as galaxy classification and gravitational lensing detection. In the domain of time-series analysis, these algorithms have significantly improved the efficiency of searching for periodic signals and identifying rare transient events.

2 Methodology and Applications

2.1 Deep Learning Architectures

The core of deep learning lies in its hierarchical structure, which allows for the automatic learning of data representations. Convolutional Neural Networks (CNNs) are particularly effective for spatial data, such as astronomical images. By using convolutional layers to capture local features, CNNs can effectively distinguish between different morphological types of galaxies. For sequential data, such as light curves or spectra, Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks are

4.4 不同红移下的模型评估

Redshift is a critical parameter in astronomical observations. Since our data has not been corrected for redshift (de-redshifted), the generalization capability of the model may be affected by redshift variations. To investigate this impact, we divided the data into several redshift intervals: $[0.0, 0.03)$, $[0.03, 0.05)$, $[0.05, 0.07)$, $[0.07, 0.10)$, $[0.10, 0.15)$, and $[0.15, 0.30)$. We then calculated the prediction error within each interval, as shown in [Figure 5: see original paper] (b).

The results indicate that within the redshift range of $[0.05, 0.15]$, the prediction error remains low and stable at approximately 0.07 dex. However, the error increases at both the high-redshift and low-redshift ends, reaching a maximum of 0.1 dex. This phenomenon can likely be attributed to two factors. First, at certain redshifts, key spectral features shift out of our sampling range; specifically, the $[OII]$ emission line is absent at specific redshifts, while at other redshifts, the $H\alpha$ and $[NII]$ emission lines also fall outside the sampled wavelength range. Second, the sample size in these extreme ranges is limited—for instance, samples with a redshift $z > 0.15$ account for only 6.12% of the total data—leading to increased prediction uncertainty.

4.5 基于 CNN 模型预测的质量-金属丰度关系

The Mass-Metallicity Relation (MZR) serves as a critical diagnostic tool for the chemical evolution of galaxies. Its accurate determination is of great significance for understanding the history of galaxy formation. In this study, we analyze this relationship by combining the gas-phase metallicities predicted by our CNN model with the stellar mass data from the MPA-JHU catalog, comparing our results with the classic study by Tremonti et al. [?]. The MZR relationship proposed by Tremonti is expressed as:

where 12+log(O/H) represents the gas-phase metallicity and M_* is the logarithm of the stellar mass relative to the solar mass ($M_\odot$). This MZR relationship is applicable over a stellar mass range of $8.5 < \log(M_*/M_\odot) < 11.5$, providing an important

reference benchmark for this study.

In this research, we utilize the test set data to analyze this relationship. The gas-phase metallicities are predicted by the CNN model, while the stellar mass data are derived from the `lgm_{fib}p50` values in the MPA-JHU catalog. We restricted the stellar mass of the test set to the range of

$8.5 < \log(M_*/M_\odot) < 11.5$ and plotted

the relationship between these variables, as shown in [Figure 6: see original paper]. The solid red line represents the fitted relationship, as shown in Equation (3):

representing the median values, while the dashed red lines indicate the

scatter of the distribution, which is approximately 0.1 dex. The black solid and dashed lines represent the median empirical MZR and its corresponding range as proposed by Tremonti et al. [?]. The results demonstrate that the MZR predicted by the CNN is consistent with the empirical MZR across most of the mass range. This further validates the accuracy and reliability of our model.

The horizontal axis represents the stellar mass ($\log(M_*/M_\odot)$) from `lgm_{fib}p50`, and the vertical axis represents the gas-phase metallicity predicted by the CNN.

6 Mass-metallicity relation: the x-axis represents stellar

mass (lgm_{fib}p50) from the MPA-JHU catalog, and the y-

axis represents the gas-phase metallicity predicted by the

CNN model.

5 应用

This chapter utilizes the trained models to predict the gas-phase metallicities of star-forming galaxies observed in LAMOST DR10 that were not included in the MPA-JHU catalog. We begin with approximately 91,400 spectra that failed to match with the MPA-JHU catalog in Chapter 2. First, we perform identification of star-forming galaxies within this dataset (see Section 5.1), and subsequently employ the trained CNN regression model to predict their gas-phase metallicities (see Section 5.2).

5.1 恒星形成星系的识别

Since the LAMOST data does not directly provide classification information for star-forming galaxies, this study constructs a binary classification model using a convolutional neural network (CNN) to distinguish between star-forming and non-star-forming galaxies. The architecture of this model is similar to the regression model described in Section 3.1 (see [Figure 2: see original paper]), with the exception that the output layer is modified from one node to two nodes, corresponding to star-forming and non-star-forming galaxies, respectively, and utilizes a Sigmoid activation function.

In [Figure 6: see original paper], the red line represents the gas-phase metallicity predicted by the CNN.

The training labels for this model are derived from the SDSS MPA-JHU catalog.

<0:02>0:21(cid:11)>0:1512+lg(O/H)=(cid:0)1:492+1:847(cid:2)lg(M(cid:3)/M)(cid:0)0:08026(cid:2)[lg(M(cid 27)$(cid : 6)\pm 1(cid : 27)9.29.08.88.68.48.28.59.09.510.010.511.0 Z pred lg(M* / M\odot) Predicted Median Line Predicted Upper 1$ $ LinePredicted Lower 1σ LineMZR Median LineMZR Upper 1σ LineMZR Lower 1σ Line3.02.52.01.51.00.5kernel density (a.u.)

Feng Jiabao et al.: Deep Learning-Based Estimation of Gas-Phase Metallicities in LAMOST Star-Forming Galaxies

The `bptclass` parameter from the DR8 catalog [?] was utilized to identify galaxy types, where `bptclass=1` corresponds to star-forming galaxies. To train the model, we randomly selected 10,000 star-forming galaxies (`bptclass=1`) and 10,000 non-star-forming galaxies from the approximately 151,000 spectra cross-matched with the MPA-JHU catalog as described in Section 2, resulting in a total sample of 20,000 spectra. These spectra underwent the same spectral interpolation and normalization procedures detailed in Section 2. Subsequently, the dataset was partitioned into training, validation, and testing sets using a ratio of 7:1:2. The training process followed a methodology similar to the regression model described in Section 3.2, employing the Binary Cross-Entropy Loss (BCELoss) function. We used the Adam optimizer with a learning rate of 0.001 and a batch size of 128, training the model for a total of 100 epochs.

The performance of the classification model was evaluated using Precision, Recall, and the F1-score. The results demonstrate that the precision, recall, and F1-scores for both star-forming and non-star-forming galaxies exceed 92%. The overall classification accuracy of the model is approximately 92.5%. Among the 7.5% of misclassified samples, approximately 87% are "Composite" galaxies. The spectral characteristics of these galaxies lie between those of star-forming galaxies and Active Galactic Nuclei (AGN); their emission line properties are particularly similar to those of star-forming galaxies, making them prone to misclassification and serving as the primary source of confusion for the model. Overall, this classification model effectively distinguishes star-forming galaxies from other types, providing a reliable foundation for constructing a high-quality sample set for subsequent regression analysis.

and non-star-forming galaxies

Galaxy Type

Precision Recall F1-Score

Star-Forming Galaxy

92.93% 92.09% 92.50%

Non-Star-Forming Galaxy

92.25% 93.07% 92.66%

We applied this classification model to approximately 91,400 spectra observed by LAMOST that were not included in the MPA-JHU catalog. To further enhance the purity of the subsequent regression sample, we implemented a confidence threshold during the inference stage, retaining only spectra with a classification confidence of as star-forming galaxies. Ultimately, the model identified approximately 20,000 high-confidence star-forming galaxy spectra. these spectral data serve as the foundation for the subsequent prediction of gas-phase metallicities.

5.2 气体金属丰度的预测与星表构建

Using the CNN regression model trained in Chapter 3, we predicted the gas-phase metallicities for the approximately 20,000 star-forming galaxies mentioned previously to construct the final catalog. presents a sample of the data from this catalog. In this table, column (1) provides the `obsid`, which serves as the unique identifier for the LAMOST spectra; columns (2) and (3) list the Right Ascension (ra) and Declination (dec) of the targets, respectively; column (4) provides the signal-to-noise ratio in the r-band (S/N_r); column (5) records the spectroscopic redshift (z); and column (6) contains the gas-phase metallicities predicted by our CNN regression model. [Figure 7: see original paper] illustrates the distribution of gas-phase metallicities within the catalog. The distribution peaks near 9.0 dex, with the majority of samples falling between 8.5 and 9.2 dex and a mean value of approximately 8.91 dex.

star-forming galaxies in LAMOST DR10

obsid

S/N_r

101162 331.90851 −1.31349

102052 331.50241 −1.30848

102075 331.56058 −1.47138

107037 332.74158 −1.25447

Note: The data in columns (1)-(5) are from the publicly available LAMOST catalog; column (6) lists the gas-

phase metallicity predicted by the model.

7 Distribution of gas-phase metallicity in our catalog

2https://pytorch.org/docs/stable/generated/torch.nn.BCELoss.html

0:85zZpredzZpred3.02.51.52.01.00.58.08.28.48.68.89.09.20DensityZpred

Acta Astronomica Sinica

Abstract

In recent years, the rapid development of deep learning has led to its widespread application across various fields of astronomical research. This paper provides a comprehensive review of the current status and progress of deep learning applications in astronomy, focusing on several key areas: galaxy morphology classification, celestial object detection and identification, astronomical image processing, and the estimation of physical parameters. We analyze the advantages of deep learning models, such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Generative Adversarial Networks (GANs), in processing large-scale astronomical survey data. Furthermore, we discuss the challenges currently faced by these methods, including the interpretability of models, the requirement for high-quality labeled datasets, and the generalization capabilities of algorithms across different observational facilities. Finally, we provide an outlook on the future trends of deep learning in the era of Big Data astronomy, particularly in the context of upcoming large-scale survey projects.

1. Introduction

With the commencement of various large-scale astronomical survey projects, such as the Large Synoptic Survey Telescope (LSST), the Square Kilometre Array (SKA), and the Wide Field Survey Telescope (WFST), astronomy has officially entered the era of Big Data. The massive volume, high dimensionality, and complexity of the data generated by these facilities pose significant challenges to traditional data processing and analysis methods. Consequently, there is an urgent need for automated, efficient, and accurate data mining techniques to extract scientifically valuable information from these vast datasets.

Machine learning, particularly deep learning—a subset characterized by multi-layered neural networks—has demonstrated exceptional performance in fields such as computer vision, natural language processing, and speech recognition. In the past decade, astronomers have increasingly adopted deep learning architectures to address complex tasks that were previously labor-intensive or computationally prohibitive. Unlike traditional methods that rely on manual feature engineering, deep learning can automatically learn hierarchical representations directly from raw data, making it uniquely suited for the nuances of astronomical imaging and spectroscopy.

2. Applications of Deep Learning in Astronomy

2.1 Galaxy Morphology Classification

The morphological classification of galaxies is fundamental to understanding galaxy formation and evolution. Traditional methods, such as visual inspection by experts or crowdsourced projects like Galaxy Zoo, are no longer feasible for the billions of galaxies expected in future surveys. Deep learning, specifically Convolutional Neural Networks (CNNs), has revolutionized this field.

Researchers have successfully applied various CNN architectures (e.g., AlexNet,

6 总结

This study utilizes LAMOST galaxy spectra to predict the gas-phase metallicity of star-forming galaxies based on a Convolutional Neural Network (CNN) model. We first performed data cleaning and filtering on the LAMOST DR10 Low-Resolution Spectroscopic (LRS) galaxy spectra, which served as the foundation for applying the CNN model to predict gas-phase metallicity. The experimental results demonstrate that the CNN model can accurately predict gas-phase metallicity, providing an efficient and precise alternative to traditional spectroscopic analysis. This approach significantly enhances data processing capabilities and prediction efficiency within a big data environment.

To evaluate the performance of the model, this study conducted assessments from multiple perspectives.

1. The precision of the model was evaluated by analyzing the deviation between the CNN predicted values and the ground truth. The results indicate that the bias between the ground truth and the values predicted by the CNN model is only 0.0034 dex, with a standard deviation of 0.0829 dex. This demonstrates that the model can accurately estimate gas-phase metallicity in the majority of cases.

2．将 CNN 与两种经典机器学习模型（随机森

We compared the Convolutional Neural Network (CNN) with other machine learning models, specifically Random Forest and Multi-Layer Perceptron (MLP). The results demonstrate that the CNN outperforms both the Random Forest and MLP models in terms of both bias and standard deviation. Notably, the CNN achieved the lowest standard deviation (0.0829 dex), indicating that this method exhibits superior performance and higher precision in the task of predicting gas-phase metallicities.

3．分析信噪比对模型性能的影响．尽管在低

Under conditions with a signal-to-noise ratio of $S/N_r = 3$, the prediction error increases to approximately 0.11 dex. Despite this increase, the model continues to demonstrate robust predictive performance across a broad range of signal-to-noise ratios.

4. We evaluated the model' s predictive performance across different redshift ranges. In the high-redshift and low-redshift regions, the prediction error increased to a maximum of approximately 0.10 dex. This variation may be attributed to the uneven distribution of spectral data or the fact that certain emission lines fall outside the sampling range; however, the overall error remains within a reasonable margin.

5．探讨模型再现星系质量-金属丰度关系方面

To evaluate the model' s performance, we performed a polynomial fit on the gas-phase metallicities predicted by the model to derive a new mass-metallicity relation (MZR). This derived relation remains fundamentally consistent with established empirical relations, demonstrating that the model can accurately recover the mass-metallicity relation of galaxies.

Finally, we applied the trained CNN model to the LAMOST DR10 LRS galaxy spectra to construct a gas-phase metallicity catalog for star-forming galaxies. First, a classification model was utilized to identify star-forming galaxies, followed by the application of the trained regression model to

predict their gas-phase metallicities. Ultimately, we obtained a catalog containing predicted gas-phase metallicities for approximately 20,000 star-forming galaxies. This provides reliable metallicity measurements for the LAMOST DR10 LRS data and offers significant data support for further astronomical research. In the future, we plan to build upon the current metallicity prediction framework by developing a multi-output or multi-task deep neural network architecture. By utilizing full-spectrum information as input, such a model will be capable of simultaneously predicting multiple galactic physical parameters (such as stellar mass and star formation rate), thereby providing a more comprehensive catalog of physical parameters for spectroscopic surveys like LAMOST.

References

Nagamine K, Reddy N, Daddi E, et al. SSRv, 2016,

202: 79

Simons R C, Papovich C, Momcheva I, et al. ApJ,

2021, 923: 203

Maiolino R, Mannucci F. A&ARv, 2019, 27: 3

Garnett D R. AJ, 1992, 103: 1330

Pérez-Montero E. MNRAS, 2014, 441: 2663

Izotov Y I, Thuan T X. ApJ, 1999, 511: 639

Pagel B E J, Edmunds M G, Blackwell D E, et al.
MNRAS, 1979, 189: 95
Alloin D, Collin-Souffrin S, Joly M, et al. A&A, 1979,
78: 200
Dopita M A, Kewley L J, Heisler C A, et al. ApJ, 2000,
542: 224
Pérez-Montero E, Contini T. MNRAS, 2009, 398: 949
Kewley L J, Dopita M A. ApJS, 2002, 142: 35
DenicolóG, Terlevich R, Terlevich E. MNRAS, 2002,
330: 69
Tremonti C A, Heckman T M, Kauffmann G, et al.
ApJ, 2004, 613: 898
Charlot S, Longhetti M. MNRAS, 2001, 323: 887
Hoyle B, Rau M M, Bonnett C, et al. MNRAS, 2015,
450: 305
Wang L L, Yang G J, Zhang J L, et al. MNRAS, 2024,
527: 10557
Ho I T. MNRAS, 2019, 485: 3569
Teimoorinia H, Jalilkhany M, Scudder J M, et al.
MNRAS, 2021, 503: 1082
Cui X Q, Zhao Y H, Chu Y Q, et al. RAA, 2012, 12:
Zhao G, Zhao Y H, Chu Y Q, et al. RAA, 2012, 12: 723
Brinchmann J, Charlot S, White S D M, et al. MNRAS,
2004, 351: 1151
(cid:0)0:0002

Estimation of Gas-Phase Metallicity in LAMOST Star-Forming Galaxies Based on Deep Learning

Abstract

Gas-phase metallicity is a fundamental physical parameter for characterizing the evolutionary state of galaxies. Traditional methods for estimating gas-phase metallicity primarily rely on the measurement of specific emission line fluxes.

However, for large-scale spectroscopic surveys like the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST), many spectra suffer from low signal-to-noise ratios (SNR) or incomplete coverage of key emission lines, which limits the efficiency and accuracy of traditional methods. In this study, we propose a deep learning-based approach to estimate the gas-phase metallicity ($12 + \log(\mathrm{O/H})$) of star-forming galaxies directly from LAMOST spectra. We utilize a 1D Convolutional Neural Network (CNN) architecture to extract features from flux-calibrated spectra. The model is trained and validated using a cross-matched sample of LAMOST DR8 and SDSS DR12, where the ground truth labels are derived from the MPA-JHU catalog. Our results demonstrate that the deep learning model achieves high precision, with a mean absolute error (MAE) of approximately 0.03 dex and a root mean square error (RMSE) of 0.05 dex. Furthermore, the model exhibits strong robustness against varying SNR levels and can provide reliable metallicity estimates even when certain diagnostic lines are weak or missing. This work highlights the potential of deep learning in processing massive spectroscopic data and provides a new tool for studying the chemical evolution of galaxies in the LAMOST survey.

---

1. Introduction

The gas-phase metallicity of a galaxy, typically represented by the oxygen abundance ($12 + \log(\mathrm{O/H})$), serves as a crucial indicator of the chemical enrichment history and the interplay between star formation, gas inflows, and outflows [**?**]. Understanding the distribution and evolution of metallicity is essential for constraining galaxy formation models.

Traditionally, gas-phase metallicity is determined through "direct methods"using temperature-sensitive lines like [O III] \$ \$4363, or through "strong-line ratios" such as R_{23}, N2, and O3N2 when direct methods are unfeasible due to the weakness of auroral lines [**?**, **?**]. While these methods are well-established, they are highly dependent on the accurate subtraction of the stellar

Estimating Gas-phase Metallicity of Star-forming Galaxies in the LAMOST Spectral Survey Using Deep Learning

FENG Jia-bao1,2 WANG Li-li2 LI Ping1,2 (1 School of Information and Control Engineering, Jilin University of Chemical Technology, Jilin 132022) (2 School of Computer and Information, Dezhou University, Dezhou 253023)

Abstract

Gas-phase metallicity is a key parameter for measuring the chemical evolution of star- forming galaxies. Accurate estimation of gas-phase metallicity is crucial for a deeper understanding of galaxy formation and evolution processes. Traditional gas-phase metallicity estimation methods rely on emission line intensity calculations, which involve complex data processing and are difficult to scale to large spectroscopic surveys. In this study, we propose a deep learning model based on a convolutional neural network (CNN) that uses the full spectrum observed by the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST) as input. The model enables automatic estimation of gas-phase metallicity in star-forming galaxies without explicit redshift correction or emission line measurement.

The CNN model consists of 8 1D convolutional layers, 4 max-pooling layers, and 1 fully connected layer, and is trained to learn the nonlinear mapping between spectral features and gas-phase metallicity values through a regression framework. Experimental results show that the model achieves a prediction error of 0.0829 dex, which is basically consistent with traditional methods. Further evaluation shows that the CNN model performs robustly across different signal-to-noise ratios and redshift ranges, and also effectively recovers the mass-metallicity relation. Finally, the trained model is applied to the LAMOST Data Release 10 Low-Resolution Survey, generating a catalog of predicted gas-phase metallicity for star- forming galaxies, which includes about 20000 galaxy spectra. The catalog is publicly available through the Science Data Bank (https://www.scidb.cn/s/UVBRzm).

Key words galaxies: gas-phase metallicity, methods: data analysis, methods: statistical, methods: deep learning