WEP Warwick Evans Publishing

International Journal of Computer Science and Information Technology

ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 1, Number 1, Year 2023
DOI: https://doi.org/10.62051/ijcsit.v1n1.22
Journal homepage: https://wepub.org/index.php/IJCSIT/index

Parameter Estimation Model for Stellar Spectral Data based on RDN

Yuwei Chen, Yadong Wu *, Weihan Zhang

[1] Sichuan University of Science and Engineering, Yibin, China
*Corresponding Author: Yadong Wu

ABSTRACT

The temperature, gravity and other key parameters of stars change during their evolution, and finding the relationship between these parameters and the life cycle of stars has always been the research focus in astrophysics. How to estimate the parameters of massive spectral data more accurately is of great significance for studying the properties of stars. However, the current conventional parameter estimation methods have problems such as exploding gradient, vanishing gradient and mismatch of newly added stellar parameters in the face of explosive growth of astronomical data, resulting in low parameter estimation accuracy. The Residual Dense neural network (RDN) model proposed in this paper is mainly improved based on two more advanced neural networks, ResNet and DenseNet. The core of RDN is a new Residual Dense Block (RDB), which includes two modules: residual and dense. The purpose of the residual module is to learn the residual between input and output and add the identity map to it, which aims to solve the problem of vanishing gradient and exploding gradient in deep network training. The dense module is where each layer is directly connected to all the layers before it, allowing better utilization of gradients and feature reuse. Its main purpose is feature extraction. The proposed model was trained on the preprocessed LAMOST DR7 dataset, making uncertainty predictions for 17 stellar parameters in LAMOST DR7 spectra with a signal-to-noise ratio (SNR) equal to or greater than 10. The results show that the proposed model has high estimation accuracy and solves the problems existing in previous methods. Compared with ResNet, DenseNet and StarNet, the key indicators such as the mean absolute error of RDN are optimized.

KEYWORDS

Spectral Data; Parameter Estimation; RDN.

1. INTRODUCTION

Stellar spectra encapsulate a wealth of physical information about distant celestial objects.The acquisition of a substantial amount of optical spectra from various astronomical bodies is integral to addressing numerous cutting-edge questions in astronomy and astrophysics.In the context of stellar evolution,parameters such as temperature,gravity,and elemental abundance undergo continuous changes.Consequently,through the analysis of stellar parameters,astronomers can gain insights into the evolutionary phases and ages of stars—features not directly discernible from spectroscopic data. Moreover,these parameters reciprocally influence stellar evolution.By scrutinizing these factors,astronomers can predict the trajectory of a star's evolution,thereby substantiating the validity of stellar theories.Consequently,the study of stellar parameters holds paramount significance for astronomers investigating the evolutionary pathways of stars.

LAMOST (Large Sky Area Multi-Object Fiber Spectroscopy Telescope),a potent tool for extensive astronomical studies,primarily focuses on large-scale celestial spectroscopic surveys to collect vast data.Through the analysis of these data,researchers delve into inquiries related to stellar properties,galaxy evolution,and dark matter.In October 2018,LAMOST initiated the second phase of its survey plan,LAMOST II,encompassing both low and medium-resolution spectroscopic surveys.

By March 2020,the first release,and subsequently in October 2021,the version 2.0 of the DR7 were made available,comprising high-quality spectra (S/N$\geqslant$10) totaling 11.43 million.The DR7 release also includes a catalog of approximately 7 million stellar spectra parameters, featuring, for some stars, the first-time inclusion of metal abundance parameters for 15 chemical elements like carbon, magnesium,and calcium.This catalog stands as the world's largest stellar parameter dataset. Nevertheless,astronomical spectra data from distant space encounter substantial noise interference during transmission,including atmospheric conditions,stray light,cosmic rays,and additional noise disturbances during reception caused by instrument instability and systematic errors.

Methods for estimating stellar physical parameters based on spectroscopy have been extensively studied,aiming to extract features from stellar spectra and mitigate the adverse effects of noise.Currently,these methods can be categorized into three main types: line index methods,template matching,and statistical estimation approaches.

The strength of line index methods lies in their strong physical significance.However,their drawback arises from the challenges in accurately extracting and precisely describing spectral lines in observed spectra,where factors such as noise and distortions come into play.These difficulties often hinder the robustness and accuracy of such methods.

Template matching is another widely embraced method in the astronomical community.For instance, field of iron abundance parameter estimation,Li et al. [1] applied the LSP3 method (LAMOST Stellar Parameter Measurement Package) to investigate the estimation of [α/Fe] in low-resolution stellar spectra obtained by LAMOST.Additionally,Xiang et al.[2] explored the application of multivariate linear regression and Kernel Principal Component Analysis (KPCA) methods in [α/Fe] parameter estimation.However,this method has its limitations: the precision of estimation relies on the template library,often generated through a specific physical model.The physical models used for spectroscopic formation are still evolving,and existing models may have distinct applicable ranges.Consequently,the limitation results in significant differences between spectroscopic libraries formed according to physical models and observed spectra,leading to suboptimal performance of corresponding parameter estimation methods in certain situations.

With the advent of the era of artificial intelligence and big data,researchers are exploring statistical estimation methods for stellar parameter estimation.This approach typically involves establishing a nonlinear mapping relationship between spectral data and physical parameters,seeking the optimal regression model through optimization methods.Statistical estimation methods serve as black-box algorithms,offering ease of implementation and applicability in situations where a comprehensive understanding of spectral regression models might be lacking.

Neural networks represent the most widely applied algorithm in the realm of statistical estimation for stellar spectroscopic parameter estimation.The introduction of neural networks into the field of astronomical stellar spectroscopic parameter estimation dates back to 1997 when Bailer-Jones et al.[3] proposed parameter estimation method based on artificial neural networks .Subsequently,numerous applications have emerged utilizing neural networks for parameter estimation.For instance,in the domains of stellar atmospheric parameters and elemental abundance,Mantega et al.[4] employed the backpropagation neural network method to estimate atmospheric physical parameters for stars with low signal-to-noise ratios.Williamsen et al.[5] utilized a feedforward neural network method for estimating atmospheric physical parameters in medium-resolution spectra.Leung et al. [6] employed Bayesian neural networks for parameter estimation of multi-element abundances,while Ting et al. [7] used a designed Payne model for parameter fitting.Nepal et al. [8] trained a convolutional neural

network to estimate atmospheric parameters and Li element abundance for approximately 40,000 stars.StarNet,proposed by Bialek,Fabbro et al. [9,10],is applied to assess the performance of synthetic stellar spectra.Zhong et al.[11] utilized a neural network matching regression algorithm to establish a nonlinear regression model for determining atmospheric physical parameters.Zhang et al. [12] introduced the SLAM model for estimating stellar parameter labels,and Xiang et al. [13] employed the SCDD model for estimating stellar atmospheric parameters.

In the field of redshift parameter estimation for stellar spectra,Hoyle et al. [14] employed deep neural networks for estimating the redshift parameters of stars.D'Isanto et al. [15],Pasquet et al.[16],Mu et al. [17],Razim et al. [18] utilized enhanced convolutional neural networks for the estimation of redshift in stellar spectra.However,the debugging of neural networks is very complex,and the setting of its parameters mainly depends on the experience of the debugger,which is difficult to ensure that the constructed network structure is optimal.Neural networks are also prone to problems such as vanishing gradient and exploding gradient,and most of the above models use simple fully connected neural networks or convolutional neural networks to establish mapping relationships.

This study addresses the enhancement of spectral depth feature extraction by introducing the Residual Dense Connection Neural Network (RDN).This network is a modification built upon the novel network structures ResNet proposed by He et al. [19] in 2016 and DenseNet introduced by Huang et al. [20] in 2017.It combines the advantages of ResNet in preserving more features while deepening the network and enhancing network propagation efficiency.The goal is to address the problem of network degradation associated with increasing network depth.Additionally,it leverages DenseNet's ability to improve feature utilization without deepening the network,treating shallow layer network outputs as inputs for deep layer networks.The RDN aims to establish an optimal regression model,effectively avoiding the aforementioned issues.

This paper is structured as follows: Section 2 covers data presentation and preprocessing methods,Section 3 elucidates the principles of the RDN model; The results of model training and evaluation are presented in Section 4; and Section 5 predominantly focuses on summarization and prospects for future research.

2. DATA

The primary dataset utilized in this study is formed by cross-matching the high-resolution spectra from LAMOST DR7 and its reference labels (mainly involving stellar parameters such as temperature (Teff),gravity (g),and abundances of 15 chemical elements,as detailed in Table 1) with the data from the APOGEE-Payne DR17 catalog.Stellar spectra exhibit highly imbalanced characteristics in parameter space,and a larger standard dataset contributes to establishing models with improved parameter estimation accuracy.However,an excessively large standard dataset may introduce challenges such as increased computational costs,prolonged processing times,and heightened noise and irrelevant data.Thus,it is essential to strike a balance between the advantages of dataset size and the associated challenges,opting for an appropriately sized standard dataset to ensure data quality,manageability,and effectiveness.

Furthermore,the transmission process introduces considerable noise,resulting in numerous outliers.Notably,data with temperatures (Teff) above 6500K and below 4000K are exceptionally scarce and exhibit irregular distribution,significantly impacting the precision of model parameter estimation.Consequently,this study excluded values with excessively large outliers,quality flags not indicating "good," and Teff values outside the range [4000,6500]K.From the approximately 2.4 million spectra data with a signal-to-noise ratio (S/N) greater than 10 in LAMOST DR7,over 100,000 data points were selected for constructing the standard group.To facilitate optimization of deep learning models,preprocessing was performed on the reference spectra before training the parameter

estimation model.The preprocessing steps primarily involved wavelength calibration,spectral resampling,and spectral normalization.

Table 1. The LAMOST DR7 medium resolution spectral reference tag used in this paper.

Number	Parameter Name	Description
1	Teff[K]	Stellar effective temperature
2	Log g	Stellar surface gravity
3	CH	The element abundance of carbon
4	NH	The element abundance of nitrogen
5	OH	The element abundance of oxygen
6	MgH	The element abundance of magnesium
7	AlH	The element abundance of aluminum
8	SiH	The element abundance of silicon
9	SH	The element abundance of sulphur
10	KH	The element abundance of potassium
11	CaH	The element abundance of calcium
12	TiH	The element abundance of titanium
13	CrH	The element abundance of chromium
14	MnH	The element abundance of manganese
15	FeH	The element abundance of iron
16	NiH	The abundance of nickel
17	CuH	The abundance of copper
18	Quality_flag	Evaluation index of spectral quality

Firstly,wavelength calibration is implemented to ensure the accuracy of spectral data by aligning the wavelengths.This calibration is crucial for precisely matching spectral features and ensuring the observed wavelengths correspond accurately to their respective physical properties.It facilitates the accurate analysis and interpretation of data.In stellar spectroscopic data,radial velocity (RV) may broaden or narrow spectral lines in the observed spectrum,thereby affecting the precision of spectral parameter estimation.Hence,in this study,wavelength calibration is conducted using the radial velocity information provided in the LAMOST catalog to address these influences on spectral data.refer with Eq1:

$$\lambda' = \frac{\lambda}{1 + \mathrm{RV}/c} \tag{1}$$

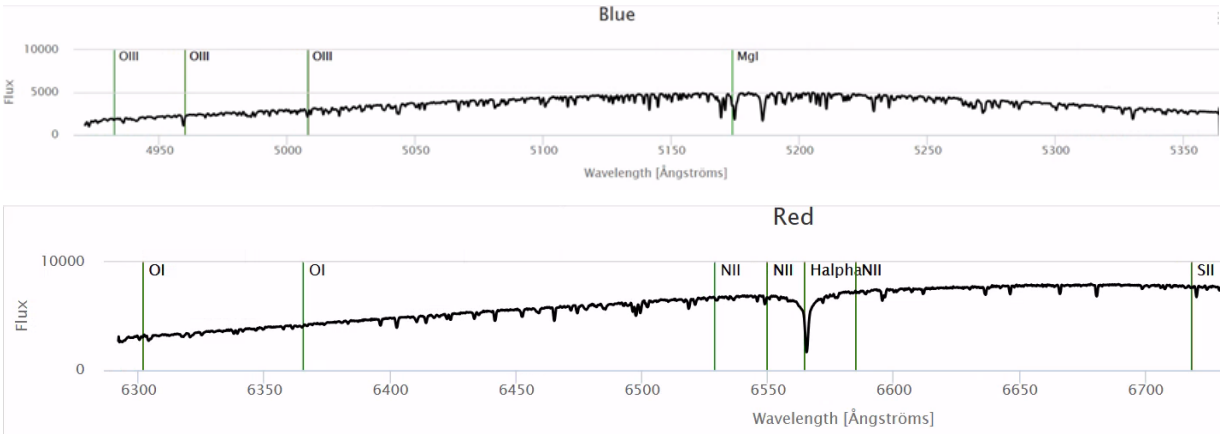

Figure 1. Example of LAMOST MRS Spectrum: Relative fluxes for the blue part(top) and red (bottom) part

Next is spectral resampling,a necessary step after wavelength calibration,as it induces changes in the wavelength distribution of spectral data.To address this,the spectra undergo resampling.In this study,an interpolation method with a step size of 0.1Å is employed for spectral resampling.As illustrated in Figure 1,the original spectral data is interpolated within the wavelength range of 4920-5360Å (blue end) and 6290-6730Å (red end).This interpolation facilitates the creation of two branches with uniform inputs for the neural network.

Finally,spectral data normalization is performed.The presence of outlier samples can prolong training times and even lead to model convergence issues.Hence,normalization of the spectral flux data is essential before training to mitigate the adverse effects of outlier samples,expedite model training,and suppress the unfavorable impacts arising from imperfect spectral flux calibration.The normalization formula is as follows:

$$\left[\widehat{X_i} = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}}\right] \tag{2}$$

Following the aforementioned preprocessing steps,the spectral data can be directly fed into the RDN model for the estimation of spectral parameters.

3. MODEL

This study employs the Residual Dense Neural Network (RDN),and Figure 2 illustrates the schematic structure of RDN,comprising four main components: shallow feature extraction,residual dense blocks,global feature extraction,and uncertainty prediction.

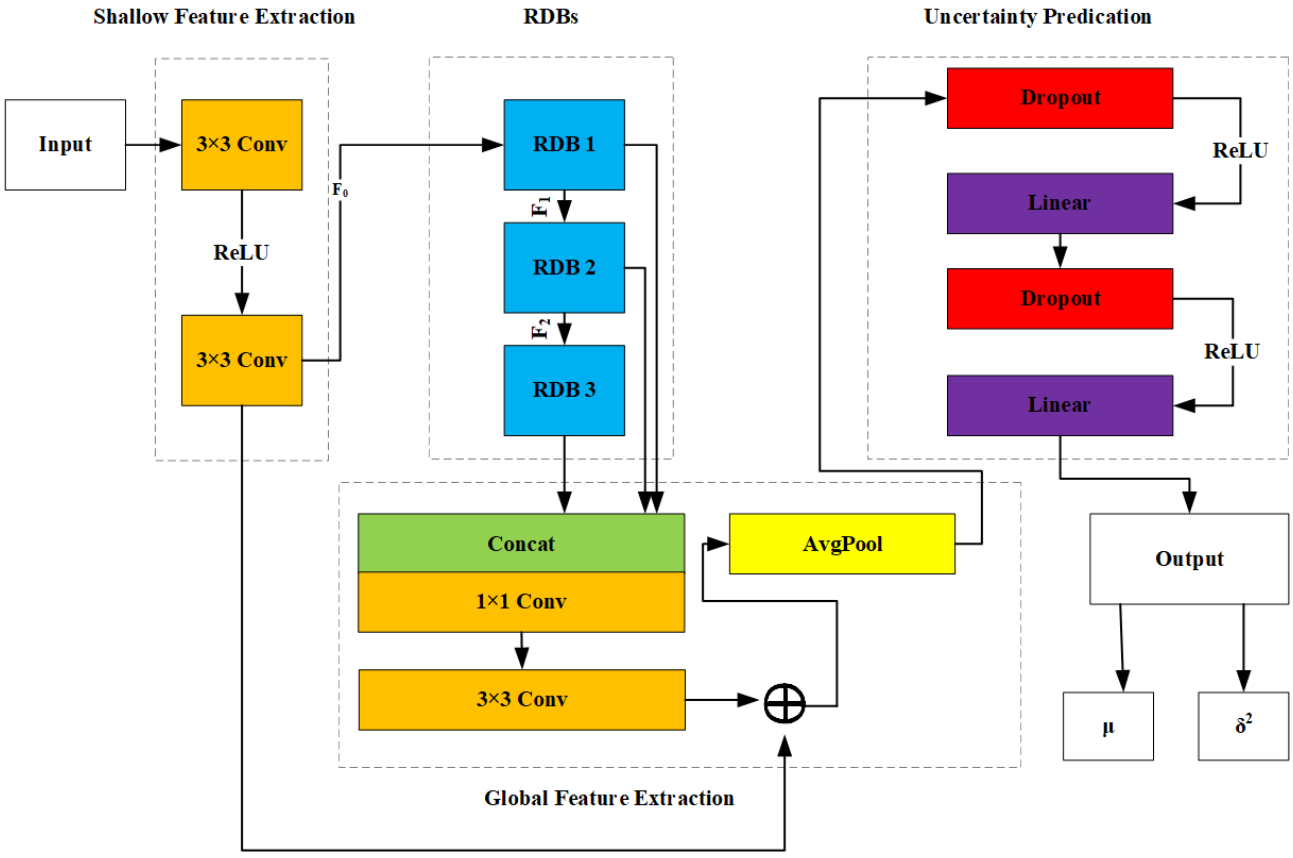

Figure 2. A diagram of the RDN model

The input consists of pre-processed spectral data,and subsequent processing involves the mentioned shallow feature extraction,residual dense blocks（RDBs）,global feature extraction,and uncertainty

prediction.Following these steps,a mapping from spectral features to probability density functions (PDF) for stellar temperature,gravity,and elemental abundance parameters is established.

The shallow feature extraction section primarily involves two convolution operations designed to extract shallow features.This process aids in extracting relevant information from the input data,enabling the network to learn and represent fundamental features conducive to subsequent higher-level representations,ultimately enhancing the overall performance of the model.

The key component demonstrating network performance is the residual dense blocks,as depicted in Figure 3.Each Residual Dense Block (RDB) consists of three concatenated residual dense units,where each residual dense unit is formed by sequentially connecting 1×1 convolution kernels and 3×3 convolution kernels containing ReLU.The green and blue solid lines in the figure represent the dense connections between each residual dense unit within the residual dense block.To enhance the connection of information between adjacent residual dense blocks,a dense connection relationship is established between the previous residual dense block and the internal residual dense units of the current block.This can reinforce the network's memory retention capability for input-output information,as indicated by the red solid line in the figure.However,an excess of dense connections may introduce redundant features by incorporating all the features from the previous residual dense block into the residual dense units of the next block,leading to issues such as excessive memory consumption and reduced accuracy.To address these problems and reduce the number of features,this study adds a 1×1 convolution at the end of each residual dense block to perform feature reduction and control the output information.

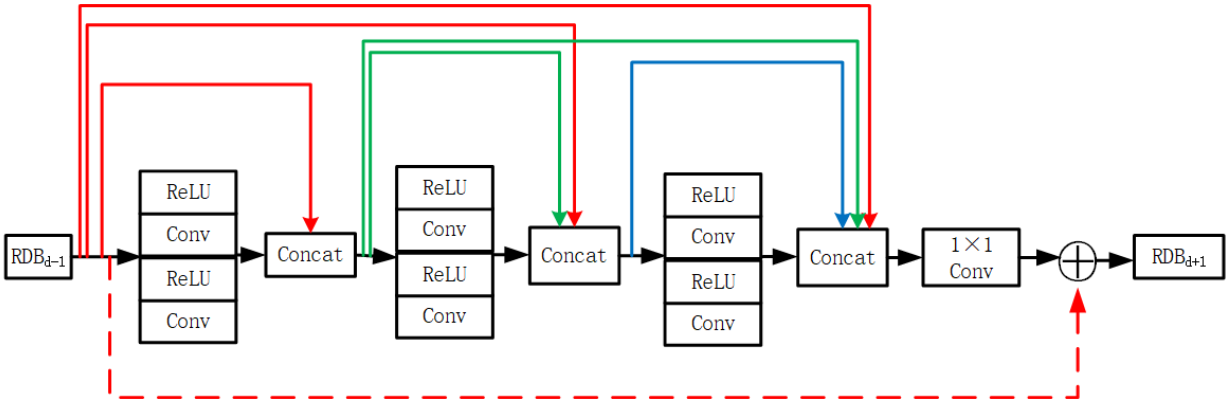

Figure 3. Schematic diagram of residual dense units in residual dense blocks(RDBs)

This study employs local residual learning,wherein the residual dense units within the residual dense block are treated as a stack of convolutional layers in a residual dense connection network.The output from the previous residual dense block serves as the input for the current module.The residual connection method is applied to add the output of the previous module to the current module's output,represented in the figure by red dashed lines and an adder.The primary purpose of local residual learning is to enhance information flow,network expressive power,and gradient transfer efficiency,facilitating rapid convergence of the network.

The global feature extraction section comprises two components: global feature fusion and global residual learning.In global feature fusion,all features extracted by the residual dense blocks are fused through concatenation and 1×1 convolution to obtain global features.Subsequently,a 3×3 convolution is applied for further feature extraction in global residual learning.Finally,a pooling layer is added to reduce network parameters and computational load,accelerating training speed and lowering model complexity.This also effectively prevents overfitting.Given the uneven distribution of stellar spectral data,average pooling is chosen over max pooling for its superior performance.Global residual learning,similar to local residual learning mentioned earlier,aims to integrate shallow features with the features extracted by all residual dense blocks,ultimately obtaining global features.

Finally,this study utilizes three fully connected layers to predict the probability density functions (PDF) for crucial parameters such as stellar temperature,gravity,and elemental abundance.The PDF are constructed using the mean (μ) and variance (δ^2) output by RDN,approximating a Gaussian distribution with mean μ and variance δ^2.Additionally,given the numerous parameters and substantial data volume in stellar spectral data,the risk of overfitting is significant.To address this,the concept of Dropout,proposed by Hinton et al.[21],is incorporated into the uncertainty prediction module.Dropout introduces a layer to mitigate overfitting during model training by reducing the shared adaptation between neurons,particularly within the same layer,and restricting a neuron's reliance on specific other neurons.However,it's noted by Hinton that employing Dropout can significantly extend training time,albeit without apparent impact during the testing phase.Consequently,this study opts not to use Dropout during the training phase to balance computational efficiency and model performance.

4. EXPERIMENT AND ANALYSIS

4.1 Model Train

The data set is the preprocessed Lamost DR7 data set mentioned in Chapter 2,which contains about 140,000 spectral data with stellar parameters.In this paper,the data set is randomly divided into training set,validation set and test set according to the ratio of 7:1:2.The training set is used for the training of the RDN model,the validation set is used for the selection of model hyperparameters,and the test set is used for the evaluation of the model.

In applications,the loss function is often linked to the optimization problem as a learning criterion,that is,the model is solved and evaluated by minimizing the loss function.In deep learning,it is often used to estimate the parameters of the model.In this paper,the negative log-likelihood of the normal distribution is used as the loss function of the model,and its formula is as follows:

$$Loss(x|\theta) = \frac{\log \sigma_\theta^2(x)}{2} + \frac{\left(y - \mu_\theta(x)\right)^2}{2\sigma_\theta^2(x)} \tag{3}$$

In this study,each model undergoes training for 30 epochs with a batch size of 256.Figures 4 and 5 illustrate the loss function curves of the model on the training and validation sets,respectively.

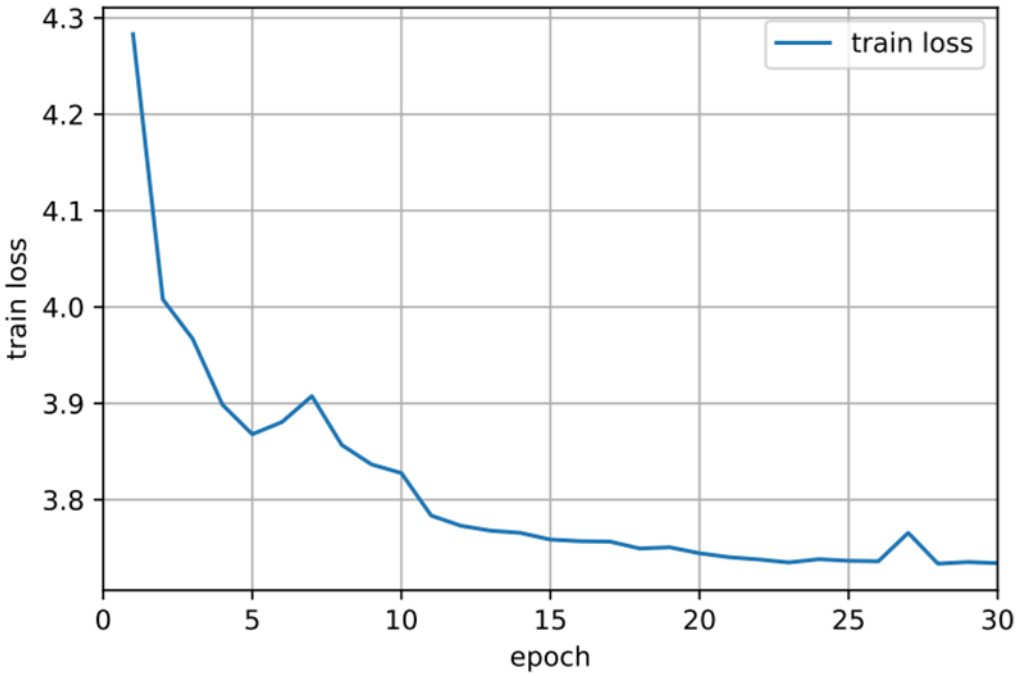

Figure 4. The relationship between epoch rounds and loss in the training set.

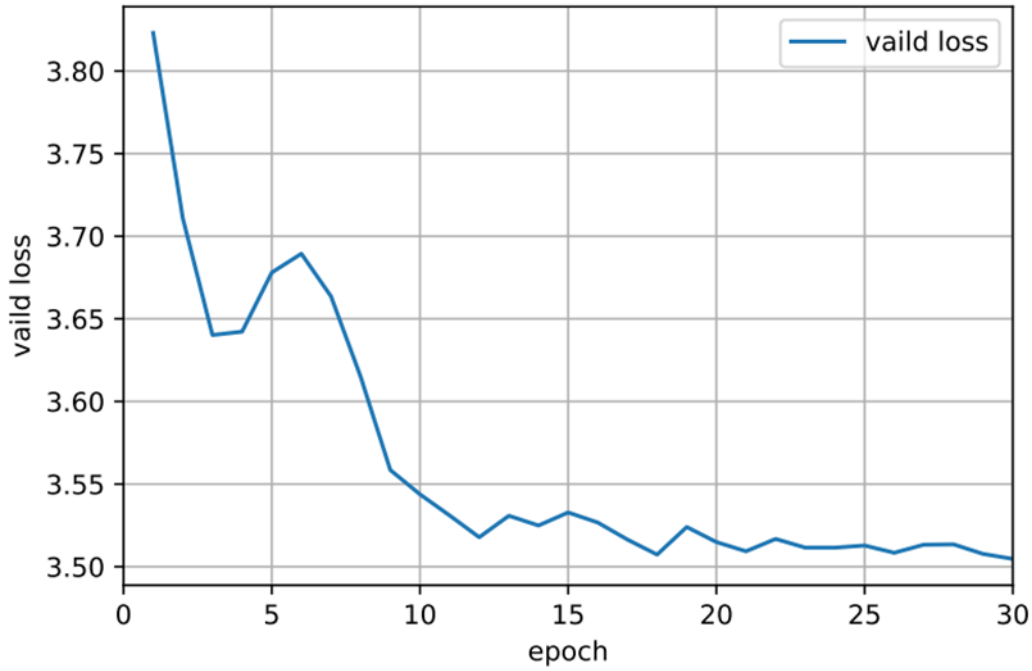

Figure 5. The relationship between epoch rounds and loss in the validation set

To enhance parameter estimation accuracy and mitigate the impact of randomness,training is conducted for M distinct instances with different random initializations (typically ranging from 5 to 7,and in alignment with Bialek et al.[9],this study adopts 6 instances).During testing,predictions from all instances are averaged to obtain the mean (μ).Subsequently,the variance (δ^2) is calculated through the computation with the mean,and the formulas for calculating μ and δ^2 are as follows:

$$\mu(y|x) = \frac{1}{M}\sum_{i=1}^{M}(y|x) \tag{4}$$

$$\sigma^2(\boldsymbol{x}) = \frac{1}{M}\sum_{i=1}^{M}\left(\sigma_{\theta_i}^2(\boldsymbol{x}) + \mu_{\theta_i}^2(\boldsymbol{x})\right) - \mu^2(\boldsymbol{x}) \tag{5}$$

In the realm of deep learning,the performance of a model is significantly influenced by the selection of its hyperparameters.The depth of the RDN model used in this study is primarily determined by the quantity of residual dense blocks (n).The Mean Absolute Error (MAE) is employed as a metric,representing the average of the absolute deviations between individual observed values and their arithmetic mean.MAE helps to mitigate issues related to the cancellation of errors,providing an accurate reflection of the actual magnitude of prediction errors.Widely utilized in statistics and deep learning,MAE serves as a metric to gauge the disparities between predicted and actual values.Evaluating the performance of the parameter estimation model using MAE offers a more nuanced assessment of the model's effectiveness,with the formula presented below:

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{6}$$

4.2 Experimental Result

Following the aforementioned training,we observed that increasing the number of residual dense blocks improves the model's performance.However,there exists a certain threshold for the number of blocks.Beyond this threshold,performance improvement becomes less pronounced and may even

regress.Taking the three crucial parameters—temperature (teff),gravity (g),and iron abundance (Fe/H)—as examples (see Table 2),it is evident that when the number of residual dense blocks is below 3,the Mean Absolute Error (MAE) demonstrates a noticeable decline.However,when the number exceeds 3,the MAE values essentially plateau,and in some cases,like temperature,there is even a slight increase.Similar trends are observed for other parameters.This phenomenon indicates that appropriately increasing the number of residual dense blocks in the model enhances the extraction capability of spectral information,thereby improving parameter estimation accuracy to a certain extent.However,an excessive number of blocks leads to increased model complexity,requiring optimization of more model parameters.When surpassing the threshold supported by the dataset and hardware,performance tends to decline.Based on these experiments,this study selects 3 as the number of residual dense blocks.

Table 2. Experimental results of different numbers of residual dense blocks (RDBs)，The model performance is measured using the Mean Absolute Error (MAE)

Number of RDBs / *Parameter*	n=1	n=2	n=3	n=4
Teff/1000	0.0561	0.0534	**0.0529**	0.0536
Log g	0.0871	0.0823	**0.0816**	0.0821
Fe/H	0.0342	0.0327	0.0300	**0.0297**

After training,the RDN model achieved precise parameter estimations: temperature (teff) at 52.92 K,gravity(log g) at 0.0816 dex,and elemental abundances (Fe, C, N, O, Mg, Al, Si, S, K, Ca, Ti, Cr, Mn, Ni, Cu) are 0.03 dex, 0.0523 dex, 0.0737 dex, 0.0750 dex, 0.0345 dex, 0.0414 dex, 0.0322 dex, 0.0572 dex, 0.0854 dex, 0.0385 dex, 0.0739 dex, 0.0645 dex, 0.0717 dex, 0.0397 dex, 0.1338 dex, ranging from 0.03 dex to 0.1338 dex.To showcase the model's superiority,we compared it with other models,such as StarNet proposed by Fabbro [9] et al.in 2017,was applied to stellar spectral studies.This paper trained StarNet using the same training set and compared the results with our model.Figure 6 shows the parameter prediction comparison of ResNet, DenseNet, StarNet, and RDN on the test set.The results indicate that RDN has lower MAE values for each parameter compared to the other three models,especially for the crucial stellar parameters, temperature, and gravity, demonstrating the significant advantage of the RDN model.

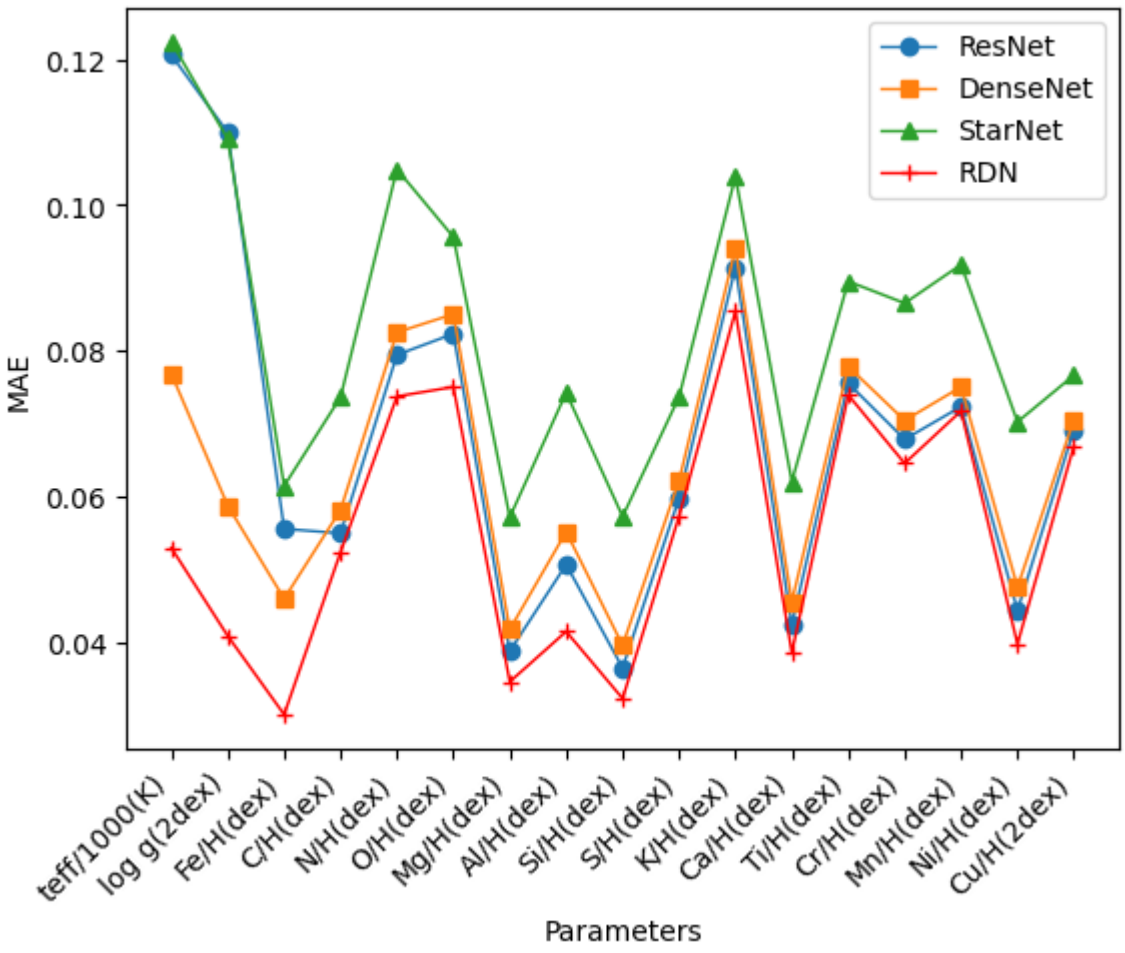

Figure 6. To standardize the orders of magnitude of MAE values for various parameters,facilitating graphical representation,this study processed selected MAE values.Specifically,the temperature was divided by 1000,the logarithm of gravity was divided by 2,and Cu/H was divided by 2.

5. CONCLUSION AND FUTURE WORK

This study introduces a Residual Dense Network (RDN) model for estimating stellar parameters.A series of preprocessing steps were applied to the Lamost DR7 dataset.Utilizing the trained RDN model,temperature,gravity,and elemental abundances were estimated for medium-resolution spectra with a signal-to-noise ratio (S/N) greater than 10 in the preprocessed Lamost DR7 dataset.Under the condition of S/N>10,the precision of temperature (teff) is 52.92K,gravity (log g) is 0.0816 dex,and the precision of 15 elemental abundances ranges from 0.03 dex to 0.1338 dex.To validate the model's performance,we compared its results with those of other models on the same dataset.The comparative analysis demonstrates the higher accuracy of the RDN model.

In the future,as various sky survey projects progress globally,both the volume of data obtained from spectral observations and the number of parameters will continue to increase.We plan to leverage the latest updated dataset from Lamost,along with datasets from GAIA,SDSS,and other releases,to integrate and retrain the model.Additionally,we aim to enhance the model,particularly in handling imbalanced data,to improve its performance for parameter estimation in a broader range of survey spectra.

REFERENCES

[1] Li J, Han C, Xiang M, et al.A method of measuring the [α/Fe] ratios from the spectra of the LAMOST survey[J].Research in Astronomy and Astrophysics,2016,16(07):86-99.

[2] Xiang M S,Liu X W,Shi J R,et al.Estimating stellar atmospheric parameters,absolute magnitudes and elemental abundances from the LAMOST spectra with Kernel-based principal component analysis[J].Monthly Notices of the Royal Astronomical Society,2017,464(3): 3657-3678.

[3] Bailer-Jones C A L,Irwin M,Gilmore G,et al.Physical parametrization of stellar spectra: the neural network approach[J].Monthly Notices of the Royal Astronomical Society,1997,292(1): 157-166.

[4] Manteiga M,Ordóñez D,Dafonte C,et al.ANNs and wavelets: A strategy for gaia RVS low S/N stellar spectra parameterization[J].Publications of the Astronomical Society of the Pacific,2010,122(891): 608.

[5] Willemsen P G,Hilker M,Kayser A,et al.Analysis of medium resolution spectra by automated methods–Application to M 55 and ω Centauri[J].Astronomy & Astrophysics,2005,436(1): 379-390.

[6] Henry W Leung,Jo Bovy,Deep learning of multi-element abundances from high-resolution spectroscopic data,Monthly Notices of the Royal Astronomical Society,Volume 483,Issue 3,March 2019,Pages 3255–3277,https://doi.org/10.1093/mnras/sty3217.

[7] Ting Y S,Conroy C,Rix H W,et al.The Payne: self-consistent ab initio fitting of stellar spectra[J].The Astrophysical Journal,2019,879(2): 69.

[8] Nepal S,Guiglion G,De Jong R S,et al.The Gaia-ESO Survey: Preparing the ground for 4MOST and WEAVE galactic surveys-Chemical evolution of lithium with machine learning[J].Astronomy & Astrophysics,2023,671: A61.

[9] Bialek S,Fabbro S,Venn K A,et al.Assessing the performance of LTE and NLTE synthetic stellar spectra in a machine learning framework[J].Monthly Notices of the Royal Astronomical Society,2020,498(3): 3817-3834.

[10] Fabbro S,Venn K A,O'Briain T,et al.An application of deep learning in the analysis of stellar spectra[J].Monthly Notices of the Royal Astronomical Society,2018,475(3): 2978-2993.

[11] Zhong-bao Liu,Juan-juan Ren,Wen-ai S,et al.Stellar spectra classification with entropy-based learning machine[J].Spectroscopy and Spectral Analysis,2018,38(2): 660-664.

[12] Zhang B,Liu C,Deng L C.Deriving the Stellar Labels of LAMOST Spectra with the Stellar Label Machine (SLAM)[J].The Astrophysical Journal Supplement Series,2020,246(1):9

[13] Xiang G,Chen J,Qiu B,et al.Estimating Stellar Atmospheric Parameters from the LAMOST DR6 Spectra with SCDD Model[J].Publications of the Astronomical Society of the Pacific,2021,133(1020):024504 (12pp).

[14] Hoyle B.Measuring photometric redshifts using galaxy images and deep neural networks[J].Astronomy and Computing,2016,16: 34-40.

[15] D'Isanto A,Polsterer K L.Photometric redshift estimation via deep learning-generalized and pre-classification-less,image based,fully probabilistic redshifts[J].Astronomy & Astrophysics,2018,609(A111): 1-16.

[16] Pasquet J,Bertin E,Treyer M,et al.Photometric redshifts from SDSS images using a convolutional neural network[J].Astronomy & Astrophysics,2019,621(A26): 1-15.

[17] Mu Y H,Qiu B,Zhang J N,et al.Photometric redshift estimation of galaxies with convolutional neural network[J].Research in Astronomy and Astrophysics,2020,20(6): 1-10.

[18] Cavuoti S,Amaro V,Brescia M,et al.METAPHOR: a machine-learning-based method for the probability density estimation of photometric redshifts[J].Monthly Notices of the Royal Astronomical Society,2017,465(2): 1959-1973.

[19] He K,Zhang X,Ren S,et al.Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition.2016: 770-778.

[20] Huang G,Liu Z,Van Der Maaten L,et al.Densely connected convolutional networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition.2017: 4700-4708.

[21] Hinton G E,Srivastava N,Krizhevsky A,et al.Improving neural networks by preventing co-adaptation of feature detectors[J].arXiv preprint arXiv:1207.0580,2012.