Astronomical Techniques and Instruments, Vol. 2, March 2025, 73–87

Review Open Access

Application of machine learning in astronomical spectral data mining

Ting Zhang[1,2], Hailong Zhang[1,2,3,4*], Yazhou Zhang[1], Xu Du[1,2], Wenna Cai[1,2], Han Wu[1,2], Yuyue Jiao[1,2], Wanqiong Wang[1], Jie Wang[1,2,4], Xinchen Ye[1,4], Jia Li[1]

[1]*Xinjiang Astronomical Observatory*, *Chinese Academy of Sciences*, *Urumqi* 830011, *China*

[2]*University of Chinese Academy of Sciences*, *Beijing* 100049, *China*

[3]*Key Laboratory of Radio Astronomy*, *Chinese Academy of Sciences*, *Nanjing* 210008, *China*

[4]*National Astronomical Data Center*, *Beijing* 100101, *China*

*Correspondence: zhanghailong@xao.ac.cn

Received: October 15, 2024; Accepted: January 3, 2025; Published Online: February 28, 2025; https://doi.org/10.61977/ati2024057; https://cstr.cn/32083.14.ati2024057

Citation: Zhang, T., Zhang, H. L., Zhang, Y. Z., et al. 2025. Application of machine learning in astronomical spectral data mining. *Astronomical Techniques and Instruments*, **2**(2): 73−87. https://doi.org/10.61977/ati2024057.

Abstract: Astronomical spectroscopy is crucial for exploring the physical properties, chemical composition, and kinematic behavior of celestial objects. With continuous advancements in observational technology, astronomical spectroscopy faces the dual challenges of rapidly expanding data volumes and relatively lagging data processing capabilities. In this context, the rise of artificial intelligence technologies offers an innovative solution to address these challenges. This paper analyzes the latest developments in the application of machine learning for astronomical spectral data mining and discusses future research directions in AI-based spectral studies. However, the application of machine learning technologies presents several challenges. The high complexity of models often comes with insufficient interpretability, complicating scientific understanding. Moreover, the large-scale computational demands place higher requirements on hardware resources, leading to a significant increase in computational costs. AI-based astronomical spectroscopy research should advance in the following key directions. First, develop efficient data augmentation techniques to enhance model generalization capabilities. Second, explore more interpretable model designs to ensure the reliability and transparency of scientific conclusions. Third, optimize computational efficiency and reduce the threshold for deep-learning applications through collaborative innovations in algorithms and hardware. Furthermore, promoting the integration of cross-band data processing is essential to achieve seamless integration and comprehensive analysis of multi-source data, providing richer, multidimensional information to uncover the mysteries of the universe.

Keywords: Machine learning; Neural networks; Stellar atmospheric parameter prediction; Stellar spectral classification

1. INTRODUCTION

Astronomical spectroscopic observations are an important part of modern astronomy, and by analyzing spectroscopic data from celestial objects, scientists can obtain key information about the chemical composition, physical properties, kinematic states, and the evolutionary history of their formation. Starting with the first detection of neutral hydrogen through the 21 cm line at 1420.4 MHz [1], the discovery of interstellar organic molecules was one of the four significant discoveries in astronomy in the 1960s, following the discovery of interstellar ammonia molecules by researchers in 1968[2] and interstellar carbon monoxide molecules in the Orion Nebula in 1970[3]. Observations of spectra have enabled astronomers to probe the molecular composition of the universe in different environments, especially in celestial bodies and celestial environments such as the interstellar medium, interstellar molecular clouds, planetary nebulae, and star-forming regions. Astronomical spectroscopy is an important tool for studying the chemical composition, physical properties, kinematics, and dynamics of objects in the universe. Spectroscopic observations can be used to study the kinematics and dynamics of objects and celestial environments in the universe[4]; for example, observations of the strength and velocity of emission lines can be used to study the motions of the rotating arms of galaxies[5], to study tidal motions of interacting galaxies, and to study turbulence and collapse in nebulae in star-forming regions[6]. Observations can be used to trace the chemical composition and evolution of objects in their regions; for example, complex molecules are often formed on dust particles in cold, dense molecular clouds[7]. Spectral observations can be used in astrophysical research. For example, the galactic plane survey of water masers can be used to study the relationship between maser distribution and star formation and evolution[8]. Different critical densities of molecules reflect the gas density of their surrounding regions; for typical Ly α emitting galaxies, the hydrogen column density is $\alpha \sim 10^{17}$–10^{20} cm^{-2}[9].

The development of astronomical observation equipment and technology has brought about a surge in the amount of observation data, and the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST)[10] is the multi-target spectroscopic telescope with the largest field of view and the highest spectral acquisition rate in the international astronomical community. The DR11 dataset released by LAMOST in September 2024 contains more than 10^8 low- and medium-resolution spectra. Similarly, the Sloan Digital Sky Survey (SDSS)[11] telescope generates approximately 10^6 spectra per observation. The Anglo-Australian Telescope (AAT), a collaboration between the United Kingdom and Australia, uses a 2-degree field (2dF) multi-fiber spectrograph[12] for its survey mission and has acquired spectra of more than one million objects. The Apache Point Observatory Galactic Evolution Experiment (APOGEE)[13] is a high-resolution and high signal-to-noise near-infrared spectroscopic survey program, and the APOGEE DR12 dataset[14] releases ~10^5 red giant spectroscopic data. The rapid increase in data volume has placed considerable pressure on subsequent data preprocessing and scientific analysis processes. Spectral data preprocessing must go through background correction, skylight background removal, wavelength calibration, flux calibration, and continuum spectra fitting[15,16]. Then, according to different scientific research objectives, the data are processed for classification, clustering, outlier analysis, and stellar atmospheric parameter measurements[17]. Scientific research on spectral data relies on accurately extracting continuous spectral data and spectral line information in the preprocessing process[18]. Traditional spectral data preprocessing and other processes rely on researchers using software such as IRAF[19] to manually process the raw spectral data using an interactive approach, and the preprocessing steps need to be repeated for each spectral data file, which consumes a large amount of manual time and results in inefficient data processing.

By carrying out processing such as classification, clustering, outlier analysis, and stellar atmospheric parameter measurements on astronomical spectra, information can be deeply mined from massive data, which is of great significance for understanding the physical properties of celestial bodies, studying the chemical compositions of celestial bodies, investigating the state of celestial motion and discovering unknown celestial bodies. With the explosive growth of spectral data, the speed of spectral data mining has failed to develop synchronously, dramatically reducing the efficiency and scientific output of spectral data processing. However, the rise of artificial intelligence technology has provided a new solution to these difficulties.

Artificial intelligence is a technological science that studies and develops theories, methods, and application systems for modeling, extending, and expanding human intelligence, often implemented using machine learning (ML) methods. ML is a subfield of artificial intelligence that studies how to enable machines to simulate or implement human learning behaviors by learning to model from data through algorithms and statistical models. Deep-learning (DL) algorithms are ML methods based on artificial neural networks, which automatically extract features layer by layer from data by constructing multi-level neural networks. Commonly used DL algorithms include convolutional neural networks (CNNs), recurrent neural networks (RNNs), and generative adversarial networks (GANs).

Computational methods were introduced to astronomy at the end of the 20th century, when simple algorithms were used to process telescope data, mainly for tasks such as orbit calculations and basic image processing. At the beginning of the 21st century, ML began to be widely used in astronomy, with algorithms such as neural networks, support vector machines, and multilayer perceptron machines being used to classify galaxies[20], photometric redshift estimation[21], and analyze cosmic microwave background data[22]. According to the data retrieved from the Astrophysics Data System (ADS), the trend of growth in the number of papers containing the terms *ML*, *DL*, or *deep neural network* in their titles, abstracts, and keywords is depicted in Fig. 1, and the application of ML in astronomy has grown significantly after 2017.

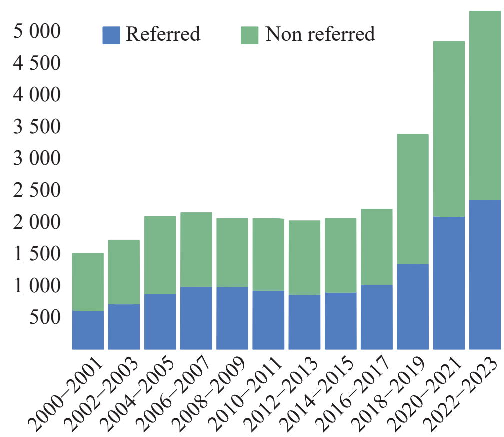

Fig. 1. Number of papers with titles, abstracts, and keywords containing ML, DL, or deep neural network in the field of astronomy, 2000–2023.

2. ML METHODS

2.1. Traditional ML Methods

Traditional ML methods include perceptron, k-nearest neighbors (kNN), decision trees, support vector machines (SVMs), clustering techniques, singular value decomposition (SVD), and principal component analysis

(PCA). The perceptron is a fundamental linear classification model and one of the earliest forms of neural networks, which performs classification by learning a linear relationship between input features and target categories. The kNN method is an instance-based, nonparametric classification and regression approach. When classifying a sample, the kNN algorithm assigns a category based on the majority vote of its kNN in the feature space. Decision trees are a model that classifies data through a series of decision nodes, where each node represents a decision rule based on a feature, branches correspond to the outcomes of these rules, and leaf nodes represent the final classification. SVM is a supervised learning algorithm primarily used for classification and regression tasks. It separates data points by constructing a hyperplane that maximizes the margin between different classes, achieving optimal classification. Clustering refers to an unsupervised learning technique to partition a dataset into distinct groups or clusters, where data points within the same cluster are highly similar, and those in different clusters exhibit significant dissimilarity. SVD is a matrix factorization method that decomposes a matrix into the product of three matrices, including singular values and singular vectors. It is widely applied for dimensionality reduction, noise filtering, and data compression. PCA is a statistical technique used for dimensionality reduction, which transforms data into a new coordinate system by identifying the principal components that capture the most variance, reducing the number of variables while retaining essential information.

Traditional ML methods generally operate based on specific assumptions or rules, resulting in relatively low model computational complexity. However, the model training process heavily relies on data preprocessing, which requires domain experts to manually extract hand-crafted features from raw data. Traditional ML methods typically require smaller datasets and have relatively low computational resource demands. Most models exhibit good interpretability because their predictions can be easily understood by analyzing model parameters or decision rules. These models usually demonstrate good generalization ability on smaller datasets but often encounter performance bottlenecks when dealing with complex, large-scale data.

2.2. DL Methods

McCulloch and Pitts[23] proposed the original artificial neural network model in 1943; Rosenblatt[24] invented the perceptual machine in 1958, regarded as the predecessor of feedforward neural networks, Rumelhart et al.[25] redeveloped the backpropagation algorithm for feedforward neural network learning in 1986, and Hinton et al.[26] introduced the concept of DL in 2006, referring to ML that includes complex neural networks such as deep neural networks.

An artificial neural network is a network-like ML model composed of neuron connections inspired by biological neural networks. It was invented to mimic the structure of the human brain and its function as an information processing system. Artificial neural networks comprise an input layer, a hidden layer, and an output layer. Fig. 2 illustrates the structure of a deep neural network with n hidden layers; the input layer receives the input data, each input node corresponds to a feature of the data, the hidden layer contains multiple nodes, which process the input data through activation functions, the number and size of the hidden layers affects the network's complexity and learning ability, and the output layer is used to generate the final output results. The output layer generates the final output, such as category labeling in a classification task or numerical prediction in a regression task.

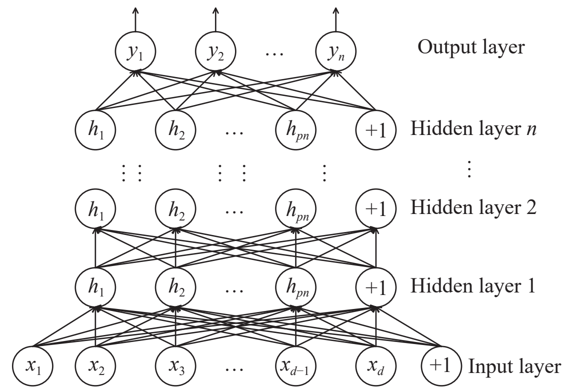

Fig. 2. Deep neural network structure with *n* hidden layers.

2.2.1. *Feedforward Neural Networks*

A feedforward neural network is the most representative neural network, composed of multiple layers of neurons, including an input layer, hidden layer, and output layer; the neurons between the layers are connected, the neurons within the layers are not connected, the output of the neurons of the former layer is the input of the neurons of the latter layer, and there is no feedback in the entire network, so the feedforward neural network cannot memorize. The input data are passed from the input layer to the output layer through the weighting and bias of each layer to produce the output result.

Artificial neural networks learn by adjusting the weights of the connections between neurons. The learning process is implemented through the backpropagation algorithm[25], which calculates the loss function to evaluate the difference between the predicted and actual values and then updates the weights through optimization algorithms such as gradient descent to make the model gradually approach the optimal solution. Feedforward neural networks are used primarily in classification problems, regression problems, pattern recognition, and other tasks.

2.2.2. *CNNs*

A CNN is a neural network for predicting image data (grid-structured data) with a hierarchical grid structure, which can be regarded as a special feedforward neural network. The basic structure of a CNN consists of an input layer, a convolutional layer, a pooling layer, a fully-connected layer, and an output layer. The convolutional layer scans the input data through a convolutional kernel, automatically extracts the low-level features such as edges and textures, and then progressively extracts the high-level features through multilayer convolution; the pooling layer is used for downsampling to reduce the size of the feature maps and reduce the complexity of the computation while retaining the key information; the fully-connected layer maps the previously extracted features to the output. The fully-connected layer maps the previously extracted features to the output, such as the results of classification or regression tasks.

The application areas of CNNs include computer vision, natural language, and speech processing. They are used in computer vision for tasks such as image classification, target detection, and image segmentation, and they are the core models in this field. Fukushima[27] proposed the Necocognitron model in 1980, and LeCun et al.[28] proposed the LeNet-5 model in 1989 based on the inverse propagation algorithm to propose the LeNet-5 model; these models achieved good recognition results on small image recognition but poor recognition results on large-scale data[29]. Then, in 2012, Krizhevsky et al.[30] developed a CNN called AlexNet to obtain the best classification results to date in the ImageNet large-scale image classification challenge competition.

The power of CNNs in processing high-dimensional data such as images and signals makes it equally suitable for spectral analysis and data classification tasks in astronomy. Shi et al.[31] designed SCNet network based on the CNN model used to classify stellar spectral images. The model was compared with many typical classification networks in DL to achieve the 0.861 highest classification accuracy; Aniyan et al.[32] used CNN for radio galaxy classification, the model architecture is depicted in Fig. 3, the number of training samples, precision, recall, F_1 score, and test samples of the model are presented in Table 1, and the classification effect is comparable to the manual classification but much faster; Keown et al.[33] used CNN for the multi-peak spectra fitting identification and classification, which was tested on 30000 datasets consisting of noisy, single-peak, and bimodal spectra, and the classification accuracies reached 100%, 99.92% ± 0.02%, and 96.72% ± 0.18%.

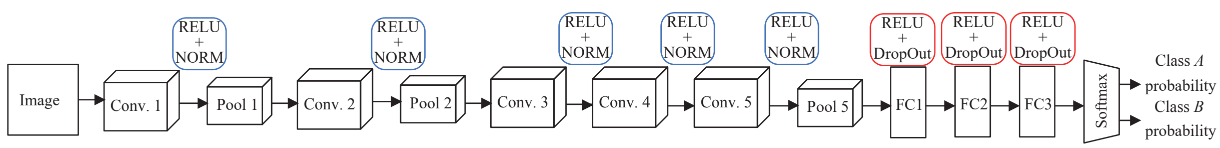

Fig. 3. CNN architecture for radio galaxy classification with output as probability scores for two galaxy classes[32].

Table 1. Class of the sources, size of the training samples for each class, precision, recall, and F_1 classification score for the validation sample and the support[32]

Class	Training samples		Precision/(%)	Recall/(%)	F_1 score/(%)	Support
	Actual	Augmented				
Bent-tailed	177	25 488	95	79	87	77
FR I	125	36 000	91	91	91	53
FR II	227	32 688	75	91	83	57
Average			88	86	86	187

2.2.3. *RNNs*

A RNN extends the traditional feedforward neural network, which can predict the input sequence data. It learns the hidden representation of the variable length input sequence through the internal recurrent hidden variables; the output of the activation function of the hidden variables at each moment depends on the output of the activation function of the recurrent hidden variables at the previous moment[34], which can use the information of the previous moment, so the RNN has the function of memory. Fig. 4 illustrates the simple architecture of the RNN.

There are various types of RNNs, including long and short-term memory networks, gated recurrent unit networks, and deep RNNs. RNNs are usually learned using backpropagation algorithms, and their application areas include natural language processing, speech processing, and time series prediction.

2.2.4. *Transformer Model*

The transformer model is a DL model based on the self-attention mechanism[35] and consists of two parts, the encoder and the decoder, each consisting of several identical sublayers stacked on top of each other. Each sublayer of the encoder mainly consists of a multi-head self-attention mechanism and a feedforward fully-connected network, which generates a weighted contextual representation by calculating the attentional weights between different positions in the input sequence and a feedforward fully-connected network that performs independent nonlinear variations of the representation at each position. Each sublayer usually consists of two fully-connected layers and an activation function. Each sublayer is followed by a residual connection, and layer normalization is used to accelerate training and stabilize the network. Each sublayer of the decoder consists of a multi-head self-atten-

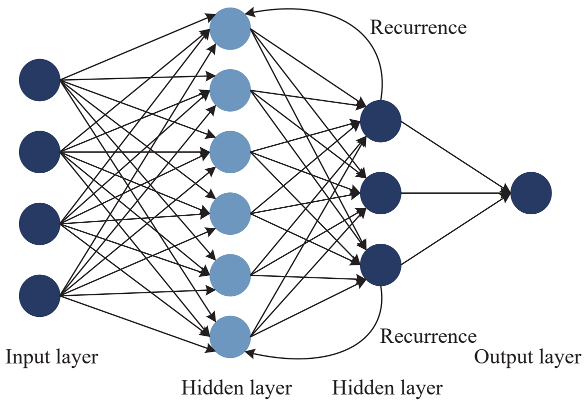

Fig. 4. RNN infrastructure.

tion mechanism, an encoder-decoder attention mechanism, and a feedforward fully-connected network, where the decoder generates a new target sequence based on the encoder's output. Because the Transformer model does not have a mechanism for sequential processing of sequences, it adds positional encoding to the input embedding to permit the use of sequential information from the input sequences. Initially used for natural language processing tasks, the Transformer model has been widely used in astronomy for processing various types of time sequences, spectral data, and image data.

As the core technology of artificial intelligence, ML has gradually become an important tool for solving astronomical data processing problems, given its powerful data processing and pattern recognition capabilities. Data classification and prediction can be conducted effectively through the statistical model in ML. Neural networks in DL, especially CNNs, have a strong adaptive ability to complex high-dimensional data. In astronomical data processing, the application of these techniques is able to not only improve the efficiency and accuracy of data processing but also provide more possibilities for exploring the hidden scientific laws of the universe. Table 2 summarizes the commonly used models and their applications of ML in astronomy in recent years[36–38].

3. APPLICATION OF ML TO SPECTRAL DATA PROCESSING

3.1. Measurements of Stellar Atmospheric Parameters

Measurement of stellar atmospheric physical parameters, including stellar surface effective temperature T_{eff}, surface gravity log g, metal abundance [Fe/H], and microscopic turbulent velocity, helps to model stars of different masses, ages, and evolutionary stages. As an important unit of galaxies, the study of the physical parameters of a large number of stellar atmospheres in galaxies can reveal the evolutionary process of galaxies. There are direct and indirect measurement methods for stellar parameters, and indirect measurement methods are the primary means of stellar atmospheric measurements at present, including photometric methods, infrared flux methods, Balmer line profile fitting, and spectral template fitting[86]. With the arrival of the astronomical big data era and the rapid development of ML algorithms, data-driven stellar atmospheric parameter measurement methods based on combining ML algorithms and spectra with large data volumes are the current major trends.

Stellar spectral data are complex and nonlinear, with intricate relationships between spectra and physical parameters. Extracting comprehensive information and understanding the relationships between these parameters are major challenges in spectral data mining[87]. With the prominence of ML in modeling nonlinear relationships, more studies have adopted ML methods to predict stellar spectral parameters. Stellar spectral parameter estimation using traditional ML methods usually includes two processes: spectral feature extraction and mapping learning. Zhang et al.[88] designed the stellar labeler SLAM based on the support vector regression technique for extracting the stellar parameters from an extensive survey spectral dataset. The SLAM model uses the cross-validation data of LAMOST DR5 and APOGEE DR15. Eight-fold cross-validation is used during training to find the best-fitting hyperparameters. The SLAM model predicts poorly for low signal-to-noise data, and the computational cost and storage requirements increase significantly as the number of labels in the training set increases. Xiang et al.[89] have used PCA for feature extraction by mapping spectra to feature space using a nonlinear function. Model training and testing revealed that the prediction results for spectral data with $R_{\mathrm{SN}} \geqslant 50$ were better than those for low SNR data.

Traditional ML methods rely on labeled data and have poor prediction results for low Signal-to-Noise Ratio (SNR) data. With the increase in data volume and the development of DL methods, artificial neural networks have been applied to predict atmospheric parameters of stellar spectra. Li et al.[90] proposed the StarGRUNet model based on artificial neural networks and self-attentive learning mechanisms. The self-attentive module was used to learn different types of spectral parameter features; the model uses data with $5 \leqslant R_{\mathrm{SN}} \leqslant 50$ and $R_{\mathrm{SN}} \geqslant 50$, the test set contains 67340 and 100973 entries each, and the test set contains 19240 and 28850 entries each. The model test results reveal that for spectral data with signal-to-noise ratios greater than or equal to 5, the prediction accuracies of StarGRUNet for T_{eff} and log g are 94 K and 0.16 dex. Li et al.[91] constructed a deep CNN based on the open-source python package of astrNN to predict the abundance of nine elements in stellar spectra simultaneously, and to avoid the prediction error caused by the uneven distribution of the sample set, the researcher added a weight matrix to the loss function. Pan et al.[56] used feedforward neural networks for stellar atmospheric parameter prediction, using 50000 spectral data released by SDSS, with training and test sets of 5000 and 45000 spectra. The results of the study indicated that the mean absolute errors for the effective temperature T_{eff}, surface

Table 2. Standard models and applications of ML in astronomy

	Standard learning methods/models	Application in astronomy
Traditional ML	Supervised learning methods: perceptual machines, kNN, Naive Bayes, logistic regression, SVMs, Boosting, decision trees, and random forests (RF)	Classification of astronomical spectra[39], Classification of active galactic nuclei[40], Classification of supernovae[41], Identification of single pulses[42], Radio-frequency interference suppression[43], Search for fast radio bursts[44], Identification of pulsar candidates[45], Classification of galactic spectra[46]
	Unsupervised learning methods: clustering methods, SVD, PCA, Markov Chain Monte Carlo	Stellar spectral analysis[47], spectral cluster analysis[48], Stellar atmospheric parameter estimation[49], Hydrogen atom clock troubleshooting[50], Galaxy spectral classification[51], Pulsar candidate identification[52], RF interference suppression[43]
DL	Feedforward neural networks	Galaxy photometric redshift estimation[53], Gamma-ray burst identification[54], Feeder compartment fusion measurement prediction[55], Stellar atmosphere parameter estimation[56]
	CNNs	Galaxy morphology classification[57], Star formation rate measurements[58], Spectral redshift estimation[59], Cosmological parameter estimation[60], Merging galaxy clusters identification[61], Galactic photometric redshift prediction[62], Coronal ejecta detection[63], Pulsar candidate identification[64], Fast radio burst classification[65], Radio-frequency interference detection[66], Gravitational wave signal detection[67], Stellar spectral classification [31], Stellar atmospheric parameter prediction[68], Fast radio burst search[69], Galaxy spectral classification[70]
	RNNs	Global 21-cm spectral line signal simulation[71], RF interference detection[72], Supernova classification[73], Strong gravitational lens parameter prediction[74], Variable star classification[75], Pulsar candidate identification[52], Stellar atmosphere parameter prediction [76], Stellar spectral classification[77]
	Generating adversarial networks	Stellar spectral classification[78], Pulsar candidate identification[79]
	Autoencoder	Gamma-ray burst identification[80], Galaxy image compression and denoising[80], Stellar spectral classification[81], Defective spectral recovery[81]
	Transformer model	Gravitational wave signal detection[82], Stellar spectral classification[83], Galaxy morphology classification[84], Photometric redshift estimation[84], Stellar atmospheric parameter prediction[85], Stellar spectral restoration [85]

gravity log g, and metal abundance [Fe/H], were respectively 79.95 K, 0.1706 dex, and 0.1294 dex, and the model uses stacked self-coding neural network to effectively overcome the problems of local minima and gradient dispersion in the training process of traditional back-propagation neural network. Wang et al.[92] proposed a stellar atmospheric parameter measurement algorithm combining CNNs and RF, using CNNs to extract features from pseudo-two-dimensional images generated based on the spectra, extracting higher-dimensional nonlinear features in the spectral data, and improving the prediction accuracy. Fabbro et al.[68] applied a deep neural network architecture to analyze the stellar spectra, using the APOGEE stellar spectral dataset to train the CNN model StarNet. StarNet extracts the spectral line feature information in the spectra by window splitting to predict the star's effective temperature, surface gravity, metal abundance, and other parameters. The StarNet model architecture is depicted in Fig. 5. The initial training of the model uses the APOGEE ASSET grid to generate 300000 synthetic spectral data, of which 224000 were used as the training set, 36000 as the validation set, and 40 000 as the test set. The StarNet model and Cannon2 model[93] were used for training and testing on 85341 spectra of the APOGEE DR12 dataset. The models were evaluated using mean absolute error (MAE) and root mean square error (RMSE) as the evaluation metrics for prediction performance. The StarNet model had superior prediction ability, as presented in Table 3. The StarNet model has superior prediction ability on high SNR spectra, which is consistent with the ASPCAP data processing standard pipeline. However, the prediction error is larger in low signal-to-noise spectra. The test reveals that the differences in the size of the training set samples, the range of stellar parameters, and the training time can cause differences in the training results.

Aiming to solve the problem of noise interference during stellar spectral feature extraction, Xiong et al.[94] proposed the residual RNN RRNet, which is used primarily for estimating stellar parameters from LAMOST medium-resolution spectra. RRNet mainly consists of residual, recurrent, and uncertainty modules and suppresses the effects of noise and irrelevant components by enhancing the spectral features. The model performance improves with the increase of hyperparameters, but when the hyperparameters reach a certain threshold, the performance improvement is weak or even decreases. Adding more residual blocks can help improve the ability to extract spectral information. However, the training set cannot be infinitely scaled to support higher complexity models in practical applications. Leung et al.[85] trained a stellar astronomical model based on the Transformer and large-scale language modeling techniques, as well as a stellar astronomi-

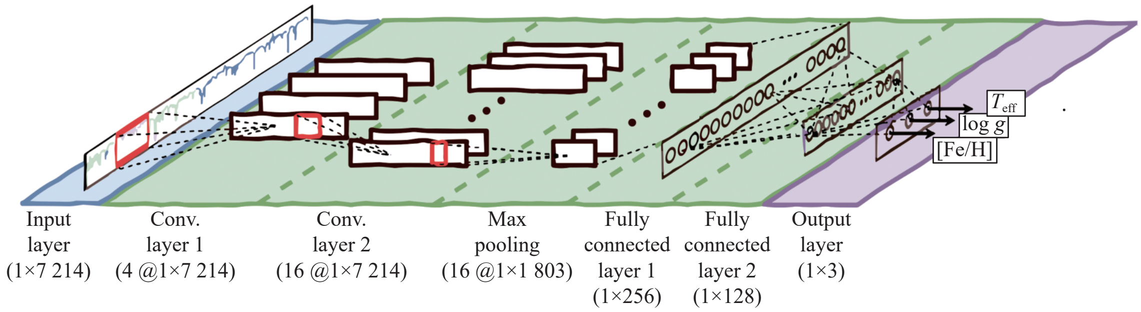

Fig. 5. Architecture of 7-layer StarNet model based on CNN[68].

Table 3. StarNetc2 model and Cannon2 model APOGEE DR12 data test results[68]

Model	Metric	$T_{\rm eff}$/K	log g/dex	[Fe/H]/dex
StarNet$_{\rm c2}$	MAE	31.2	0.053	0.025
	RMSE	51.2	0.081	0.040
Cannon2	MAE	46.8	0.066	0.036
	RMSE	71.6	0.102	0.053

cal base model. Fig. 6 illustrates the model's architecture, which performs generative and discriminative tasks, including implementing stellar parameter extraction, spectral generation and restoration, and mapping between stellar parameters. The model was trained by a self-supervised learning approach using data such as APOGEE and Gaia. In the task of mapping from spectra to stellar parameters, the model's prediction accuracies for the $T_{\rm eff}$, log g, and [M/H] parameters are 47 K, 0.11 dex, and 0.07 dex, which illustrate comparable accuracies to that of the fine-tuned XGBoost model. The model can handle multiple tasks without fine-tuning and provide prediction inaccuracy, making it competitive with traditional machine learning models.

3.2. Classification of Stellar Spectra

Traditional spectral classification methods usually rely on researchers comparing spectra with standard stellar samples by the naked eye[95], which, despite its pattern recognition and classification capabilities, is inherently subjective, and the process of visual inspection through the human eye is highly time-consuming and difficult to handle large amounts of data. With the increasing volume and complexity of astronomical data, traditional spectral classification methods are becoming obsolete, and the solution to these problems is the development of automated methods based on the quantitative determination of spectral features. von Hippel et al.[96] proposed a pattern recognition-based approach to achieve spectral classification by mimicking the operations of a human classifier when visually inspecting spectra and estimating their presence. Scibelli et al.[97] achieved spectral recognition based on searching for spectra that are most similar to the observed spectra in a library of template spectra of different spectral models, correlating the observed spectra with the spectra in the template spectral library, and classifying to which class of spectra they belong based on the magnitude of their correlation coefficients. These automated

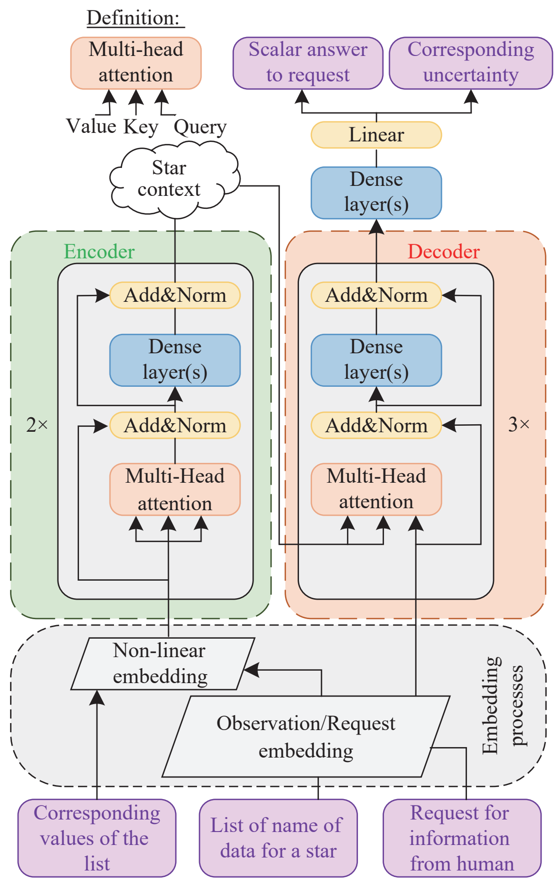

Fig. 6. Neural network architecture of the transformer-based stellar astronomical base model[85].

techniques are an improvement over naked-eye classification. However, they are difficult to apply to spectral data of different resolutions generated by different observing instruments and observing modes, and the pattern-matching method suffers from low efficiency and few standard star spectra.

With the deepening of machine learning research and applications, more machine learning methods are used to analyze and process astronomical spectral data. Liu et al.[98] applied SVMs to the classification of stellar spectra and found that the completeness of classification was as high as 90% for A- and G-type stars, but for O-, B-, and K-type stars, the completeness was as low as 50%, resulting in about 40% of the O-, B-, and K-type stars being respectively were misclassified as A- and G-type stars. Thus, traditional machine learning methods are prone to large errors when facing multi-class classification tasks, especially when the classes are unbalanced. Zhang et al.[99] establish a stellar spectral classification model based on multi-class SVMs, which solves the problem of higher complexity of SVMs when targeting multi-classification problems. However, model prediction accuracy is not high. Zhang et al.[100] propose the XGBoost-based stellar spectral feature classification method that uses the XGBoost algorithm (for automatic classification and feature ranking) to obtain the known or unknown spectral lines most sensitive to the classification decision.

Traditional machine learning algorithms can extract several layers of stellar spectral features, cannot extract high-level features, and are susceptible to the imbalance of the training set, and therefore DL methods have been gradually applied to the stellar spectral classification problem. Zheng et al.[101] proposed a spectral generation method based on a one-dimensional generative adversarial network used to balance the training sample dataset, and then a CNN was used for the classification task. The model architecture is depicted in Fig. 7, and the model is used for O-, B-, A-, F-, G-, K-, and M-type star classification, and the average correct rate of classification is 95.3%. Liu et al.[102] proposed a supervised algorithm for stellar spectral classification on the basis of one-dimensional stellar spectral CNN to classify F-, G-, and K-type stellar spectra and their subclasses and compared the model with the Artificial Neural Network Algorithms, random forest algorithms, and SVM algorithms. The one-dimensional CNN model has the highest classification accuracy on the same dataset. Hong et al.[103] used CNN to extract the deep features of spectra and combine them with the attention mechanism to learn the important spectral features. It reduced the spectral dimensions by pooling operations to compress the number of model parameters. The method achieved an accuracy of 92.04% in the classification of F-, G-, and K-type stellar spectra. Wang et al.[81] developed a deep neural network-based automatic astronomical spectral feature extraction method applied to astronomical spectral classification and defective spectral repair. The method uses a pseudo-inverse learning algorithm to train a multilayer neural network layer by layer to automatically extract features from spectral data. It uses a softmax regression model for classification. When the number of training samples is 608483, and the number of networks in the model is increased to 12, the model has F_1 scores of 0.8549, 0.7891, and 0.8499 for the classification of F-, G-, and K-type stars. Sharma et al.[104] proposed to use a CNN model for the automatic classification of stellar spectra with the model architecture depicted in Fig. 8, which uses an autoencoder for pre-training of unlabeled spectral data, adjusting the weights of the encoding and decoding layers, and then supervised training based on the labeled data.

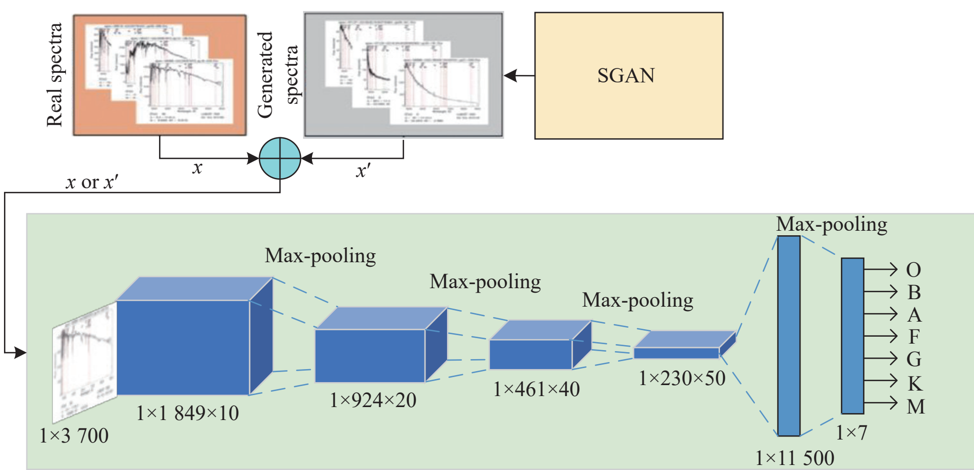

Fig. 7. Combined CNN and SGAN architecture[101].

On the basis of the remarkable progress of CNNs in astronomical spectral classification, researchers have proposed several improved models to increase classification accuracy and processing efficiency. Liu et al.[105] proposed a hybrid DL network BERT-CNN combining Transformer and CNN, which captures the intrinsic relationship between spectral features through the self-attention mechanism in Transformer, compresses the features through the pooling layer in the CNN, and finally integrates the features through the all-connectivity layer and outputs the classification results through a softmax classifier. Han et al.[106] proposed a stellar spectral classification model MSFnet based on multiscale feature fusion, using the multiscale fusion module to preprocess the data and inputting the processed data into a four-layer CNN model for the classification task, and successfully classified the stellar spectra into the six types of B-, A-, F-, G-, K-, and M-type. Zou et al.[107] proposed a convolutional network that combines the residual mechanism and attention mechanism for astronomical spectral classification, using convolutional operations to extract shallow and deep features into the spectral data, the residual mechanism to

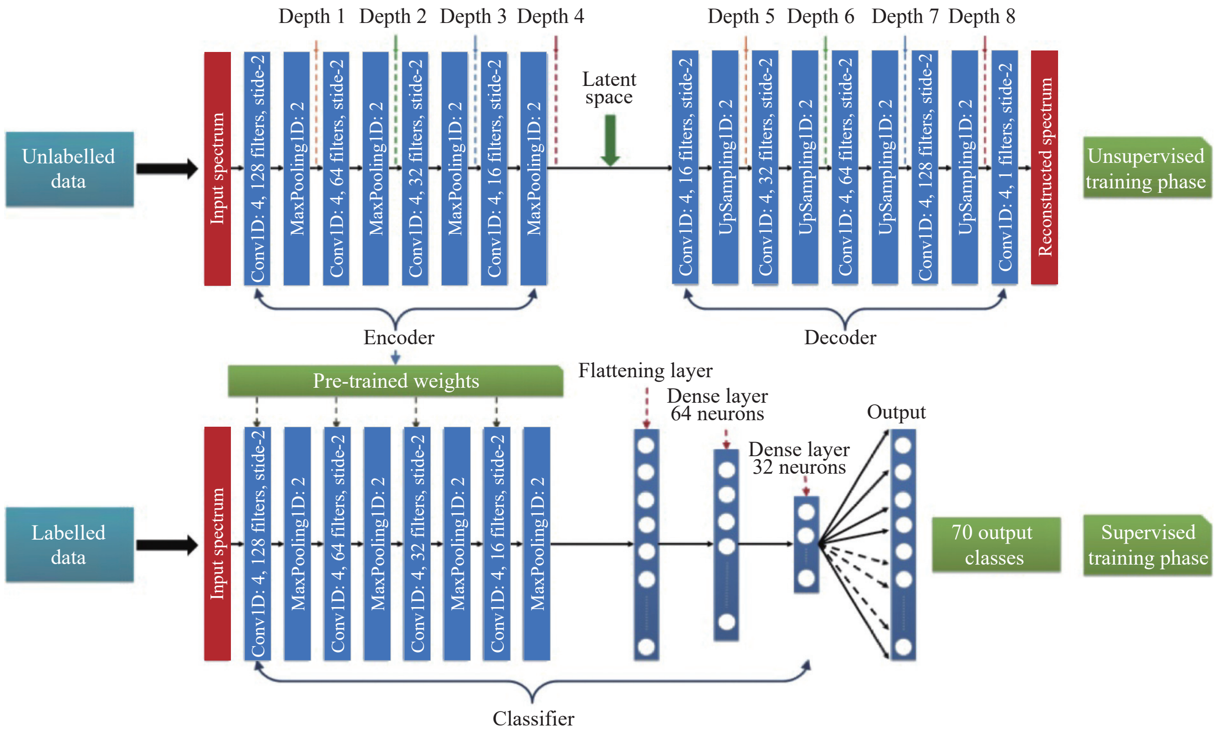

Fig. 8. Semi-supervised one-dimensional CNN classification model architecture[104].

increase the depth of the network, and the attention mechanism to enable the network to focus on specific spectral bands and features to make the learning process more targeted. Liu et al.[78] proposed StellarGAN, a stellar spectral classification model based on GANs, and the training process of the model is depicted in Fig. 9; in the pre-training stage, the model is trained using unlabeled stellar spectra and spectra generated by generative networks, and the fine-tuning phase is retrained using labeled spectra. The StellarGAN model was trained with SVM, RF, PCA, LLE, and CNN using the same number of training and test sets, where the StellarGAN model obtained the highest F_1 score (0.63) on the SDSS dataset using only 1% of the labeled data, compared with SVM, RF, CNN and other traditional methods, StellarGAN has a clear advantage in small sample learning. The experimental results also reveal a direct relationship between the spectral signal-to-noise ratio and the F_1 score. When the signal-to-noise ratio is low, the F_1 score also decreases.

4. CONCLUSIONS

This paper systematically analyzes the application of machine learning methods in stellar spectral atmospheric parameter prediction and stellar spectral classification, where Table 4 summarises the test results of the relevant literature in the task of predicting stellar spectral atmospheric parameters, and Table 5 shows the test results of the relevant literature in the task of classifying stellar spectral. In the stellar spectral atmospheric parameter prediction task, CNN and Transformer models perform the best in prediction. Although the CNN model performs well on small-scale data, Transformer is more advantageous for large-scale data. However, the performance of all the mod-

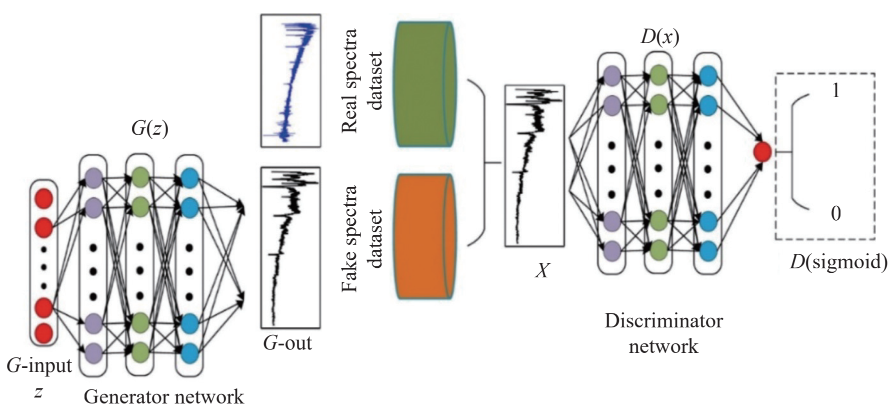

Fig. 9. Schematic of the StellarGAN model training process[78].

Table 4. Literature test results related to the application of ML in the prediction of stellar spectral atmospheric parameters

	Model	Training set	Test set	Indicators	T_{eff}/K	log g/dex	[Fe/H]/dex	[M/H]/dex
Zhang et al.[88]	SVR	17175	8171443	Prediction error	49	0.10	—	0.037
Li et al.[90]	Artificial Neural Networks(ANNs)	168313	48090	MAE	49.28	0.084	0.041	—
Li et al.[91]	CNNs	68363	7596	MAE	29	0.07	0.03	—
Pan et al.[56]	FNNs	5000	45000	MAE	79.95	0.1706	0.1294	—
Fabbro et al.[68]	CNNs	12681	85341	MAE	31.2	0.053	0.025	—
Xiong et al.[94]	RNNs	80812	23198	MAE	51.7068	0.0808	0.0308	—
Leung et al.[85]	Transformer	397718	44080	Prediction error	47	0.11	—	0.07

Table 5. Literature test results related to the application of machine learning on the classification of stellar spectra

	Model	Spectral type	Training set	Test set	Accuracy	Precision	Recall	Harmonic mean F_1
Zhang et al.[100]	XGBoost	F	840375	280125	0.8988	—	—	—
		G	1507318	502440	0.8537	—	—	—
		K	357752	119251	0.9234	—	—	—
Zhang et al.[100]	GAN CNNs	F	1000	400	0.958	—	—	—
		G	1000	400	0.960	—	—	—
		K	1000	400	0.975	—	—	—
Wang et al.[81]	DNNs	F	12994	2293	—	—	—	0.8468
		G	16448	2903	—	—	—	0.7747
		K	13058	2304	—	—	—	0.8427
Sharma et al.[104]	CNNs	F	155	27	—	0.92	0.98	0.95
		G	251	44	—	0.89	0.89	0.89
		K	221	39	—	0.88	0.89	0.89
Liu et al.[105]	Transformer CNNs	F	3585	1536	0.9108	0.9565	0.9708	0.9656
		G	1779	762	0.9239	0.9620	0.9659	0.9639
		K	2269	972	0.9311	0.9696	0.9809	0.9752
Han et al.[106]	CNNs	F	5762	1921	0.951	0.961	0.952	0.956
		G	5763	1922	0.957	0.96	0.953	0.956
		K	5766	1922	0.975	0.957	0.975	0.966

els is degraded on low signal-to-noise spectral data, and the computational cost and storage requirements increase significantly with large-scale datasets. For stellar spectral classification, CNNs far outperform traditional machine learning methods. Generative multi-resistance network methods achieve better classification results by combining a large amount of unlabeled data while relying on a small amount of labeled data, and cross-modal models based on the Transformer method can achieve the same level of classification ability as supervised learning models while achieving multi-classification. Transformer-based cross-modal models can realize multi-class tasks while achieving the same classification ability as supervised learning models.

Artificial intelligence techniques still have limitations when performing astronomical spectral analysis tasks. In the classification problem, the uneven distribution of training sample data and low spectral signal-to-noise ratio affect the model performance, and part of the model performance relies on a large amount of high-quality labeled data, which requires human resources and time for noise reduction and labeling. This requirement is because astronomical data processing involves a complex physical background, the "black-box" characteristic of the model makes it difficult to generate an explanation, the consumption of computational resources is significant, and model training and reasoning require a large amount of computational resources.

Future advancements could focus on developing data augmentation techniques to create high-quality simulated datasets using generative adversarial networks. Additionally, DL approaches that remove the need for manual feature extraction from raw data could be explored, along with the adoption of semi-supervised or unsupervised learning methods. These approaches would enable the use of novel structures within unlabeled data, expanding training datasets or creating standardized datasets.

Furthermore, the development of interpretable models capable of providing scientifically meaningful insights should be prioritized. Establishing foundational models of astronomical phenomena and training generalized models using large-scale survey and multi-band astronomical datasets could significantly reduce the time needed for iterative model training and inference. With the development of computer hardware and software technology and the improvement of artificial intelligence algorithms, the combination of artificial intelligence and astronomy is an inevitable trend in the development of astronomy. Artificial intelligence-related technology can provide astronomical data analysis with higher processing efficiency of spectral line data, realize seamless docking and comprehensive analysis of multi-source data, and provide richer, multidimensional information support for the comprehensive revelation of the mysteries of the universe.

ACKNOWLEDGEMENTS

This work is supported by the National Key R&D Program of China (2021YFC2203502 and 2022YFF0711502); the National Natural Science Foundation of China (NSFC) (12173077); the Tianshan Talent Project of Xinjiang Uygur Autonomous Region (2022TSYCCX0095 and 2023TSYCCX0112); the Scientific Instrument Developing Project of the Chinese Academy of Sciences (PTYQ2022YZZD01); China National Astronomical Data Center (NADC); the Operation, Maintenance and Upgrading Fund for Astronomical Telescopes and Facility Instruments, budgeted from the Ministry of Finance of China (MOF) and administrated by the Chinese Academy of Sciences; Natural Science Foundation of Xinjiang Uygur Autonomous Region (2022D01A360).

AUTHOR CONTRIBUTIONS

Ting Zhang contributed to the conceptualization of the study, provided research support, drafted the manuscript, and conducted revisions. Hailong Zhang and Jie Wang critically reviewed the manuscript, oversaw the research process, and offered guidance on the interpretation of content and analysis of results. Yazhou Zhang, Jie Wang, Wenna Cai, Xu Du, and Xinchen Ye also reviewed the manuscript and provided additional supervisory insights. Han Wu, Yuyue Jiao, Wanqiong Wang, and Jia Li contributed research support. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

Hailong Zhang is an editorial board member for Astronomical Techniques and Instruments and was not involved in the editorial review or the decision to publish this article. The authors declare no competing interests.

REFERENCES

[1] Ewen, H. I., Purcell, E. M. 1951. Observation of a line in the galactic radio spectrum: Radiation from galactic hydrogen at 1,420 Mc/sec. *Nature*, **168**: 356.

[2] Cheung, A. C., Rank, D. M., Townes, C. H., et al. 1968. Detection of NH 3 molecules in the interstellar medium by their microwave emission. *Physical Review Letters*, **21**(25): 1701−1705.

[3] Wilson, R. W., Jefferts, K. B., Penzias, A. A. 1970. Carbon monoxide in the Orion nebula. *The Astrophysical Journal*, **161**: L43.

[4] Sormani, M. C., Treß, R. G., Glover, S. C. O., et al. 2019. The geometry of the gas surrounding the Central Molecular Zone: on the origin of localized molecular clouds with extreme velocity dispersions. *Monthly Notices of the Royal Astronomical Society*, **488**(4): 4663−4673.

[5] Li, Z., Gerhard, O., Shen, J., et al. 2016. Gas dynamics in the milky way: a low pattern speed model. *The Astrophysical Journal*, **824**(1): 13.

[6] Ostriker, E. C., Shetty, R. 2011. Maximally star-forming galactic disks I. Starburst regulation via feedback-driven turbulence. *The Astrophysical Journal*, **731**(1): 41.

[7] Bergin, E. A., Alves, J., Huard, T., et al. 2002. N_2H^+ and $C^{18}O$ depletion in a cold dark cloud. *The Astrophysical Journal*, **570**(2): L101.

[8] Walsh, A. J., Breen, S. L., Britton, T., et al. 2011. The H_2O southern galactic plane survey (HOPS)–I. Techniques and H_2O maser data. *Monthly Notices of the Royal Astronomical Society*, **416**(3): 1764−1821.

[9] Gronke, M., Bull, P., Dijkstra, M. 2015. A systematic study of Lyα transfer through outflowing shells: model parameter estimation. *The Astrophysical Journal*, **812**(2): 123.

[10] Cui, X. Q., Zhao, Y. H., Chu, Y. Q., et al. 2012. The large sky area multi-object fiber spectroscopic telescope (LAMOST). *Research in Astronomy and Astrophysics*, **12**(9): 1197.

[11] James, E. G., Walter, A. S., Edward, J. M. 2006. The 25 m telescope of the sloan digital sky survey. *The Astronomical Journal*, **131**(4): 2332−2359.

[12] Lewis, I. J., Cannon, R. D., Taylor, K., et al. 2002. The Anglo-Australian Observatory 2dF facility. *Monthly Notices of the Royal Astronomical Society*, **333**(2): 279−298.

[13] Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011. Sdss-III: Massive spectroscopic surveys of the distant universe, the milky way, and extra-solar planetary systems. *The Astronomical Journal*, **142**(3): 72.

[14] Alam, S., Albareti, F. D., Prieto, C. A., et al. 2015. The eleventh and twelfth data releases of the Sloan Digital Sky Survey: final data from SDSS-III. *The Astrophysical Journal Supplement Series*, **219**(1): 12.

[15] Weilbacher, P. M., Palsa, R., Streicher, O., et al. 2020. The data processing pipeline for the MUSE instrument. *Astronomy & Astrophysics*, **641**: A28.

[16] Song, D. Y., Zhang, J. B., Wang, L. 2023. BFOSC Long-slit spectrum automatic extraction pipeline. *Astronomical Research and Technology*, **20**(6): 591−599. (in Chinese)

[17] Yang, H. F. 2016. Astrospectral data mining and analysis. Beijing:Electronic ublishing House. (in Chinese)

[18] Luo, F., Liu, C., Zhao, Y. H. 2019. Automatic normalization method of stellar spectrum based on spline function. *Astronomical Research and Technology*, **16**(3): 300−311. (in Chinese)

[19] Tody, D. 1986. The IRAF data reduction and analysis system. In Proceedings of SPIE, 627: 733−748.

[20] Ball, N. M., Loveday, J., Fukugita, M., et al. 2004. Galaxy types in the Sloan Digital Sky Survey using supervised artificial neural networks. *Monthly Notices of the Royal Astronomical Society*, **348**(3): 1038−1046.

[21] Firth, A. E., Lahav, O., Somerville, R. S. 2003. Estimating photometric redshifts with artificial neural networks. *Monthly Notices of the Royal Astronomical Society*, **339**(4): 1195−1202.

[22] Auld, T., Bridges, M., Hobson, M. P., et al. 2007. Fast cosmological parameter estimation using neural networks. *Monthly Notices of the Royal Astronomical Society: Letters*, **376**(1): L11−L15.

[23] McCulloch, W. S., Pitts, W. 1943. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, **5**: 115−133.

[24] Rosenblatt, F. 1958. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, **65**(6): 386.

[25] Rumelhart, D. E., Hinton, G. E., Williams, R. J. 1986. Learning representations by back-propagating errors. *Nature*, **323**(6088): 533−536.

[26] Hinton, G. E., Osindero, S., Teh, Y. W. 2006. A fast learning algorithm for deep belief nets. *Neural computation*, **18**(7): 1527−1554.

[27] Fukushima, K. 1980. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological cybernetics*, **36**(4): 193−202.

[28] LeCun, Y., Bottou, L., Bengio, Y., et al. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, **86**(11): 2278−2324.

[29] Yu, K., Jia, L., Chen, Y. Q., et al. 2013. Deep learning: Yesterday, today, and tomorrow. *Joural of Computer Research and Development*, **50**(9): 1799−1804. (in Chinese)

[30] Krizhevsky, A., Sutskever, I., Hinton, G. E. 2012. Imagenet classification with deep convolutional neural networks. In Proceedings of Advances in Neural Information Processing Systems 25.

[31] Shi, J. H., Qiu, B., Luo, A. L., et al. 2023. Stellar classification with convolutional neural networks and photometric images: a new catalogue of 50 million SDSS stars without spectra. *Monthly Notices of the Royal Astronomical Society*, **520**(2): 2269−2280.

[32] Aniyan, A. K., Thorat, K. 2017. Classifying radio galaxies with the convolutional neural network. *The Astrophysical Journal Supplement Series*, **230**(2): 20.

[33] Keown, J., Di Francesco, J., Teimoorinia, H., et al. 2019. CLOVER: Convnet line-fitting of velocities in emission-line regions. *The Astrophysical Journal*, **885**(1): 32.

[34] Chung, J., Glcehre, C., Cho, K. H., et al. 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv:1412.3555*.

[35] Ren, H., Wang, X. G. 2021. Review of attention mechanism. *Journal of Computer Applications*, **2021,41**(Suppl 1): 1−6. (in Chinese)

[36] Tao, Y. H., Cui, C. Z., Zhang, Y. X., et al. 2020. Application and improvement of deep learning in astronomy. *Progress in Astronomy*, **38**(2): 168−188. (in Chinese)

[37] Zhang, H. L., Wang, J., Zhang, Y. Z., et al. 2024. Review of artificial intelligence applications in astronomical data processing. *Astronomical Techniques and Instruments*, **1**(1): 1−15.

[38] Smith, M. J., Geach, J. E. 2023. Astronomia ex machina: a history, primer and outlook on neural networks in astronomy. *Royal Society Open Science*, **10**(5): 221454.

[39] Jones, O. C., Woods, P. M., Kemper, F., et al. 2017. The SAGE-Spec Spitzer Legacy program: the life-cycle of dust and gas in the Large Magellanic Cloud. Point source classification–III. *Monthly Notices of the Royal Astronomical Society*, **470**(3): 3250−3282.

[40] Hassan, T., Mirabal, N., Contreras, J. L., et al. 2013. Gamma-ray active galactic nucleus type through machine-learning algorithms. *Monthly Notices of the Royal Astronomical Society*, **428**(1): 220−225.

[41] Toshikage, S., Tanaka, M., Yasuda, N., et al. 2024. A systematic search for rapid transients in the Subaru HSC-SSP transient survey. *The Astrophysical Journal*, **977**(1): 18.

[42] Ling, Y., Zhang, J. Q., Li, X. R., et al. 2022. Research on recognition method of single-pulse signals based on XGBoost. *Astronomical Research and Technology*, **19**(3): 264−273. (in Chinese)

[43] Zhang, H. L., Zhang, Y. Z., Wang, J., et al. 2022. Application of RFI mitigation technology in radio astronomy. *Astronomical Research and Technology*, **19**(6): 613−625. (in Chinese)

[44] Wagstaff, K. L., Tang, B., Thompson, D. R., et al. 2016. A machine learning classifier for fast radio burst detection at the VLBA. *Publications of the Astronomical Society of the Pacific*, **128**(966): 084503.

[45] Tan, C. M., Lyon, R. J., Stappers, B. W., et al. 2018. Ensemble candidate classification for the LOTAAS pulsar survey. *Monthly Notices of the Royal Astronomical Society*, **474**(4): 4571−4583.

[46] Tao, H., Zhang, Y. X., Cui, C. Z., et al. 2018. Automated spectral classification of galaxies using machine learning approach on Alibaba cloud AI platform (PAI). *arXiv:1801.04839*.

[47] Bailer-Jones, C. A. L., Irwin, M., Von Hippel, T. 1998.

Automated classification of stellar spectra—II. Two-dimensional classification with neural networks and principal components analysis. *Monthly Notices of the Royal Astronomical Society*, **298**(2): 361−377.
[48] Wang, K., Guo, P., Luo, A. L. 2015. Angular quantization based affinity propagation clustering and its application to astronomical big spectra data. In Proceedings of 2015 IEEE International Conference on Big Data (Big Data).
[49] McGurk, R. C., Kimball, A. E., Ivezić, Ž. 2010. Principal component analysis of sloan digital sky survey stellar spectra. *The Astronomical Journal*, **139**(3): 1261.
[50] Li, R. K., Li, X. R., Cai, Y. 2020. Research on fault diagnosis of hydrogen maser mased on machine learning. *Astronomical Techniques and Instruments*, **17**(3): 349−356. (in Chinese)
[51] Zhang, X., Zhang, J. N., Zhao, Y. H., et al. 2020. Spectral classification of galaxies based on clustering analysis. *Astronomical Research and Technology*, **17**(2): 233−243. (in Chinese)
[52] Wang, W. Q., Wang, J., Ye, X. C., et al. 2025. Application of AI technology in pulsar candidate identification. *Astronomical Techniques and Instruments*, **2**(1): 27−43.
[53] Zhou, X. C., Gong, Y., Meng, X. M., et al. 2022. Extracting photometric redshift from galaxy flux and image data using neural networks in the CSST survey. *Monthly Notices of the Royal Astronomical Society*, **512**(3): 4593−4603.
[54] Crupi, R., Dilillo, G., Della Casa, G., et al. 2024. Enhancing gamma-ray burst detection: evaluation of neural network background estimator and explainable AI insights. *Galaxies*, **12**(2): 12.
[55] Lu, C. M., Li, M. H., Song, B. N., et al. 2024. Research on fusion measurement prediction of FAST feed cabin based on BP neural network. *Progress in Astronomy*, **42**(3): 519−528. (in Chinese)
[56] Pan, R. Y., Li, X. R. 2016. Stellar atmospheric parameterization based on deep learning. *Acta Astronomica Sinica*, **57**(4): 379−388. (in Chinese)
[57] Dieleman, S., Willett, K. W., Dambre, J. 2015. Rotation-invariant convolutional neural networks for galaxy morphology prediction. *Monthly Notices of the Royal Astronomical Society*, **450**(2): 1441−1459.
[58] Yang, Y. J., Liu, C., Yang, M., et al. 2024. The star formation history of nearby galaxies: A machine learning approach. *The Astrophysical Journal*, **977**(1): 11.
[59] Zhou, X. C., Gong, Y., Zhang, X., et al. 2024. Accurately estimating redshifts from CSST slitless spectroscopic survey using deep learning. *The Astrophysical Journal*, **977**(1): 69.
[60] Min, Z. W., Xiao, X., Ding, J. C., et al. 2024. Deep learning for cosmological parameter inference from a dark matter halo density field. *Physical Review D*, **110**(6): 063531.
[61] Arendt, A. R., Perrott, Y. C., Contreras-Santos, A., et al. 2024. Identifying galaxy cluster mergers with deep neural networks using idealized Compton-y and X-ray maps. *Monthly Notices of the Royal Astronomical Society*, **530**(1): 20−34.
[62] Henghes, B., Thiyagalingam, J., Pettitt, C., et al. 2022. Deep learning methods for obtaining photometric redshift estimations from images. *Monthly Notices of the Royal Astronomical Society*, **512**(2): 1696−1709.
[63] Xian, X. G., Shang, Z. H., Yuan, M. Y., et al. 2022. Detection algorithm of coronal mass ejections based on faster R-CNN. *Astronomical Research and Technology*, **19**(1): 65−77. (in Chinese)
[64] Liu, Y., Jin, J., Zhao, H. Y., et al. 2023. MFPIM: A deep learning model based on multimodal fusion technology for pulsar identification. *The Astrophysical Journal*, **954**(1): 86.
[65] Agarwal, D., Aggarwal, K., Burke-Spolaor, S., et al. 2020. FETCH: A deep-learning based classifier for fast transient classification. *Monthly Notices of the Royal Astronomical Society*, **497**(2): 1661−1674.
[66] Kerrigan, J., Plante, P. L., Kohn, S., et al. 2019. Optimizing sparse RFI prediction using deep learning. *Monthly Notices of the Royal Astronomical Society*, **488**(2): 2605−2615.
[67] Krastevm, P. G. 2020. Real-time detection of gravitational waves from binary neutron stars using artificial neural networks. *Physics Letters B*, **803**: 135330.
[68] Fabbro, S., Venn, K. A., O'Briain, T., et al. 2018. An application of deep learning in the analysis of stellar spectra. *Monthly Notices of the Royal Astronomical Society*, **475**(3): 2978−2993.
[69] Zhang, Y. G., Gajjar, V., Foster, G., et al. 2018. Fast radio burst 121102 pulse detection and periodicity: a machine learning approach. *The Astrophysical Journal*, **866**(2): 149.
[70] Zhong, F. C., Napolitano, N. R., Heneka, C., et al. 2024. Galaxy spectra neural Network (GaSNet) II. Using deep learning for spectral classification and redshift predictions. *Monthly Notices of the Royal Astronomical Society*, **532**(1): 643−665.
[71] Dorigo Jones, J., Bahauddin, S. M., Rapetti, D., et al. 2024. 21CMLSTM: A fast memory-based emulator of the Global 21 cm signal with unprecedented accuracy. *The Astrophysical Journal*, **977**(1): 19.
[72] Burd, P. R., Mannheim, K., März, T., et al. 2018. Detecting radio frequency interference in radio-antenna arrays with the recurrent neural network algorithm. *Astronomische Nachrichten*, **339**(5): 358−362.
[73] Charnock, T., Moss, A. 2017. Deep recurrent neural networks for supernovae classification. *The Astrophysical Journal Letters*, **837**(2): L28.
[74] Morningstar, W. R., Hezaveh, Y. D., Levasseur, L. P., et al. 2018. Analyzing interferometric observations of strong gravitational lenses with recurrent and convolutional neural networks. *arXiv:1808.00011*.
[75] Naul, B., Bloom, J. S., Pérez, F., et al. 2018. A recurrent

neural network for classification of unevenly sampled variable stars. *Nature Astronomy*, **2**: 151−155.

[76] Corral, L. J., Fierro-Santillán, C. R., Navarro, S. G. 2023. Stellar parameter estimation in O-type stars using artificial neural networks. *Astronomy and Computing*, **45**: 100760.

[77] Corral, L. J., Fierro-Santillán, C. R. 2021. Stellar spectra models classification and parameter estimation using machine learning algorithms. *arXiv:2105.07110*.

[78] Liu, W., Cao, S., Yu, X. C., et al. 2024. StellarGAN: Classifying stellar spectra with generative adversarial networks in SDSS and APOGEE sky surveys. *The Astrophysical Journal Supplement Series*, **271**(2): 53.

[79] Yin, Q., Li, Y. F., Li, J. J., et al. 2023. Pulsar-candidate selection using a generative adversarial network and ResNeXt. *The Astrophysical Journal Supplement Series*, **264**: 2.

[80] Graff, P., Feroz, F., Hobson, M. P., et al. 2014. SKYNET: an efficient and robust neural network training tool for machine learning in astronomy. *Monthly Notices of the Royal Astronomical Society*, **441**(2): 1741−1759.

[81] Wang, K., Guo, P., Luo, A. L. 2017. A new automated spectral feature extraction method and its application in spectral classification and defective spectra recovery. *Monthly Notices of the Royal Astronomical Society*, **465**(4): 4311−4324.

[82] Zhao, T. Y., Lyu, R. X., Wang, H., et al. 2023. Space-based gravitational wave signal detection and extraction with deep neural network. *Communications Physics*, **6**: 212.

[83] Li, S. C., Tu, L. P., Li, X., et al. 2023. Besvm: A-type star spectral subtype classification algorithm based on transformer feature extraction. *Spectroscopy and Spectral Analysis*, **43**(5): 1575−1581. (in Chinese)

[84] Parker, L., Lanusse, F., Golkar, S., et al. 2024. AstroCLIP: a cross-modal foundation model for galaxies. *Monthly Notices of the Royal Astronomical Society*, **531**(4): 4990−5011.

[85] Leung, H. W., Bovy, J. 2024. Towards an astronomical foundation model for stars with a transformer-based model. *Monthly Notices of the Royal Astronomical Society*, **527**(1): 1494−1520.

[86] Yuan, H. L., Zhang, Y. X., Zhang, H. T., et al. 2018. Survey of stellar atmosphere parameter estimation. *Astronomical Research and Technology*, **15**(3): 257−265.

[87] Gao, W., Li, X. R. 2016. Application of multi-task sparse group Lasso feature extraction and support vector machine regression in the stellar atmospheric parametrization. *Acta Astronomica Sinica*, **57**(4): 389−401. (in Chinese)

[88] Zhang, B., Liu, C., Deng, L. C. 2020. Deriving the stellar labels of LAMOST spectra with the Stellar LAbel Machine (SLAM). *The Astrophysical Journal Supplement Series*, **246**(1): 9.

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

[90] Li, X. R., Lin, B. Y. 2023. Estimating stellar parameters from LAMOST low-resolution spectra. *Monthly Notices of the Royal Astronomical Society*, **521**(4): 6354−6367.

[91] Li, Z. H., Zhao, G., Chen, Y. Q., et al. 2022. The stellar parameters and elemental abundances from low-resolution spectra–I 1.2 million giants from LAMOST DR8. *Monthly Notices of the Royal Astronomical Society*, **517**(4): 4875−4891.

[92] Wang, L. L., Tu, L. P., Li, S. C. 2023. A method for measuring stellar atmospheric parameters based on CNN and RF . *Science Technology and Engineering*, **23**(31): 13464−13471. (in Chinese)

[93] Casey, A. R., Hogg, D. W., Ness, M., et al. 2016. The Cannon 2: A data-driven model of stellar spectra for detailed chemical abundance analyses. *arXiv:1603.03040*.

[94] Xiong, S. C., Li, X. R., Liao, C. X. 2022. A model RRNet for spectral information exploitation and LAMOST medium-resolution spectrum parameter estimation. *The Astrophysical Journal Supplement Series*, **261**(2): 36.

[95] Morgan, W. W., Keenan, P. C., Kellman, E. 1943. An atlas of stellar spectra, with an outline of spectral classification. Chicago: University of Chicago Press.

[96] von Hippel, T., Storrie-Lombardi, L. J., Storrie-Lombardi, M. C., et al. 1994. Automated classification of stellar spectra: Where are we now? In Proceedings of a Workshop of the Vatican Observatory.

[97] Scibelli, S., Newberg, H. J., Carlin, J. L., et al. 2014. Census of blue stars in SDSS DR8. *The Astrophysical Journal Supplement Series*, **215**(2): 24.

[98] Liu, C., Cui, W. Y., Zhang, B., et al. 2015. Spectral classification of stars based on LAMOST spectra. *Research in Astronomy and Astrophysics*, **15**(8): 1137.

[99] Zhang, J., Liu, Z. B., Song, W. A., et al. 2018. Stellar spectra classification method based on multi-class support vector machine. *Spectroscopy and Spectral Analysis*, **38**(7): 2307−2310. (in Chinese)

[100] Zhang, X., Luo, A. L. 2019. XGBOOST based stellar spectral classification and quantized feature. *Spectroscopy and Spectral Analysis*, **39**(10): 3292−3296. (in Chinese)

[101] Zheng, Z. P., Qiu, B., Luo, A. L., et al. 2020. Classification for unrecognized spectra in lamost dr6 using generalization of convolutional neural networks. *Publications of the Astronomical Society of the Pacific*, **132**(1008): 024504.

[102] Liu, W., Zhu, M., Dai, C., et al. 2019. Classification of large-scale stellar spectra based on deep convolutional neural network. *Monthly Notices of the Royal Astronomical Society*, **483**(4): 4774−4783.

[103] Hong, S. X., Zou, Z. Q., Xu, L. Z. 2021. Classification model of stellar spectral type and luminosity type based on convolution neural network. *Acta Astronomica Sinica*, **62**(5): 16−28. (in Chinese)

[104] Sharma, K., Kembhavi, A., Kembhavi, A., et al. 2020. Application of convolutional neural networks for stellar spectral classification. *Monthly Notices of the Royal Astronomical Society*, **491**(2): 2280−2300.

[105] Liu, Z. B., Wang, J. 2022. Research on the improvement of spectral classification performance with the high-performance hybrid deep learning network. *Spectroscopy and Spectral Analysis*, **42**(3): 699−703. (in Chinese)

[106] Han, B. C., Song, Y. H., Zhao, Y. H. 2024. Classification of star spectrum based on multi-scale feature fusion. *Spectroscopy and Spectral Analysis*, **44**(8): 2284−2288. (in Chinese)

[107] Zou, Z. Q., Zhu, T. C., Xu, L. Z., et al. 2020. Celestial spectra classification network based on residual and attention mechanisms. *Publications of the Astronomical Society of the Pacific*, **132**(1010): 044503.