APPLICATION OF ARTIFICIAL INTELLIGENCE (AI) TECHNOLOGIES TO ASTROPHYSICAL DATA

Fayzullayeva Saodatxon Qakhramon qizi

Fergana State University, Faculty of Physics and Mathematics, Department of Physics.

Abstract. Artificial intelligence is an interdisciplinary research field with widespread applications. It aims at developing theoretical, methodological, technological, and applied systems that simulate, enhance, and assist human intelligence. Recently, notable accomplishments of artificial intelligence technology have been achieved in astronomical data processing, establishing this technology as central to numerous astronomical research areas such as radio astronomy, stellar and galactic (Milky Way) studies, exoplanets surveys, cosmology, and solar physics. This article systematically reviews representative applications of artificial intelligence technology to astronomical data processing, with comprehensive description of specific cases: pulsar candidate identification, fast radio burst detection, gravitational wave detection, spectral classification, and radio frequency interference mitigation.

Furthermore, it discusses possible future applications to provide perspectives for astronomical research in the artificial intelligence era.

1.INTRODUCTION

Astronomy is an observation-based science, for the development of which large volumes of data are essential. Recently, advances in manufacturing and information technology have considerably enhanced the data-collection capabilities of astronomical observation instruments, introducing the "big data" methodology into astronomy. With the development of an increasing number of large telescopes, large-scale digital sky surveys have become prominent in astronomy.

In recent years, China has established several large astronomical ground-based observation facilities, such as the Five-hundred-meter Aperture Spherical Telescope (FAST) [1] , the Shanghai 65 m Radio Telescope (Tianma) [2] , and the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST) [3] . Other large telescopes, for example, the Qitai 110 m Radio Telescope (QTT) [4] , are also under construction. Moreover, China has launched the Dark Matter Particle Explorer (DAMPE or Wukong) [5] and Advanced Space-based Observatory-Solar (ASO-S) [6] satellites, with plans to launch and operate other space telescopes in succession. Future large astronomical instruments will generate daily considerable vol-umes of observation data. For example, the current data volume generated by a single sky survey represents tens of terabytes and survey dataset size is gradually increasing [7] . This considerable data volume increase requires new processing methods, such as the big data methodology, creating new, specific challenges. Astronomy has become a data-driven and data-intensive science [8] . Realtime processing, transmission, storage, and analysis of large data volumes are serious problems that must be addressed.

AI is a branch of computer science that aims to study and develop theories, methods, technologies, and applied systems capable of simulating, supporting, and extending human intelligence. It establishes principles to implement computer systems that mimic human intelligence, enabling it to perform higher-level applications. As an emerging technology, AI has found widespread applications in various fields [9]. Following the "third industrial revolution" represented by the emergence of the internet and mobile communication technology, the

emergence of AI, jointly with big data, is seen as initiating the "fourth industrial revolution". The AI framework encompasses three standard paradigms for data structure extraction:

(1) Supervised learning [10] : data consist of pairs of input items and labels. Supervised learning is considered a particular powerful form of nonlinear regression that extracts patterns from existing item/label pairs to predict labels for new items.

(2) Unsupervised learning [11]: input data lack labels; therefore, the objective is to identify underlying statistical structures. Unsupervised learning is an extension of traditional statistical techniques such as clustering and principal component analysis (PCA). Modern AI systems also often rely on self-supervised learning, which achieves results similar to unsupervised learning by automatically labeling data, for example, by applying the same label to artificially generated variations of an object.

(3) Reinforcement learning [12]: its purpose is to establish methods or strategies to achieve specific objectives, using feedback from previous actions.

Astronomy includes several interconnected fields, such as astrophysics and astrochemistry. Its purpose is to identify and formulate the physical laws that govern the properties and evolution of the universe by observing astronomical phenomena, such as celestial body motion or cosmic redshift [13] . Astronomical research relies on extensive data and computational power. For specialized tasks, AI technology often outperforms traditional methods that rely on human intervention or rule-based programming; therefore, it often achieves comparable or better performance than human scientists [14] . One of the most frequent applications for AI in astronomy is image processing. With AIbased image analysis and processing techniques, astronomers discover celestial objects and analyze phenomena more rapidly. For example, during systematic galactic surveys, AI image-processing techniques allow fast and accurate assessment of galactic motion and discrimination between evolution patterns of distinct galaxies. Therefore, this technology has become essential for astronomy and has improved human understanding of the universe [15] . Astronomical research requires large volumes of data, especially for continuous observations. AI is a useful tool for data processing, providing astronomers with fast and accurate handling of extensive datasets, with the means to extract valuable information from them. This is essential to investigate the physical laws of the universe and to analyze the internal structures of celestial bodies [16] . For example, for astrophysics research, particle accelerators are vital tools that simulate cosmic conditions, such as magnetic fields and particle flows [17] . Such simulations require substantial computational power.   AI technology assists astronomers by optimizing computational resources and enhancing efficiency, thereby saving time and financial resources.

The application of AI technology in astronomy originated in the 1990s [18] , when artificial neural networks first yielded promising experimental results [19] . As indicated by statistical data from "arXiv:astro-ph" (search on papers with titles, abstracts, or keywords containing terms or abbreviations such as "machine learning", "ML", "artificial intelligence", "AI", "deep learning", or "neural networks"; https://arxiv.org/) [16] and from the astrophysics paper database (search on "machine learning" and "deep learning"; https://ui.adsabs.harvard.edu/) [20] , AI applications for astronomy emerged in the early 2000s, with an increasing interest after approximately 2015. Collectively, these papers indicate that AI technology has been extensively applied in diverse astronomical research fields, often with better performance than traditional methods. In particular, its effectiveness was demonstrated for the analysis and processing of extremely large astronomical datasets (in a bigdata context).

This article primarily provides a review of representative applications of AI technology in the field of astronomical data processing and presents the latest developments in this domain.

Additionally, it discusses future applications of AI in astronomy, providing new perspectives for astronomical data processing. The remaining sections of this article are organized as follows: Section 2 primarily introduces the application of AI technology to pulsar search; Section 3 focuses on its role for fast radio burst (FRB) detection; Section 4 describes AI application for gravitational wave (GW) detection; Section 5 outlines the use of AI technology for spectrum classification; Section 6 discusses its role for radio frequency interference (RFI) mitigation; Section 7 includes a summary of the article and details the perspectives for AI development in astronomical data processing.

2.PULSAR CANDIDATE IDENTIFICA-TION

Pulsars are extremely dense celestial objects with powerful magnetic fields, formed in the late evolutionary stages of stars with masses between 8 and 25 times that of the Sun, often after a supernova explosion. The central body of a pulsar is typically a collapsed neutron star with a material density billions of times greater than that of water [21] . Radiation from pulsars emanates from the magnetic polar regions of the neutron star. When these align with the Earth's line of sight, the emitted radiation can be detected. By analyzing pulsar radiation signals, the distribution and properties of the interstellar matter and magnetic field in the Milky Way can be studied [22] . In standard pulsar search methods, the research protocol is divided into numerous consecutive steps, including several data preprocessing tasks such as "dispersion" and matched filtering. After processing, a list of potential pulsar candidates is obtained, from which human judgment identifies true pulsar signals. This approach requires considerable data storage space and computational resources, resulting in a large number of pulsar candidates and low search efficiency [23] . The introduction of AI technology has markedly improved the efficiency of pulsar searches and candidate identification, allowing for more comprehensive analyses of the spatial distribution and temporal evolution of pulsars. Morello et al. introduced a neural-network-based pulsar identification method called Straightforward Pulsar Identification using Neural Networks (SPINN, Fig. 1) [24] , designed to manage the increasing data volumes collected during pulsar searches. This method was cross-validated on candidate pulsars identified during the High Time Resolution Universe (HTRU) survey. It demonstrated its ability to identify all known pulsars present in the survey data, while maintaining an extremely low false-positive rate of 0.64%. The SPINN method reduces the number of pulsar candidates by four orders of magnitude; thus, it will be applicable to future large-scale pulsar searches.

To improve the speed and accuracy of single-pulse searches, Pang et al. proposed a new two-stage, singlepulse search approach that combined unsupervised and supervised machine learning to automatically identify and classify single pulses from radio pulsar search data [25]. Results showed that this method detected 60 known pulsars from a benchmark dataset and 32 additional known pulsars not included in the benchmark dataset. Five faint single-pulse candidates emitting multiple periodic signals were also detected and warrant further investigation.

Additionally, Devine et al. introduced a machinelearning-based method to identify and classify dispersed pulse groups (DPGs) in single-pulse search outputs [26]. This method comprises two stages. First, peak identification is conducted with the Recursive Algorithm for Peak IDentification (RAPID). RAPID identifies local maxima by detrending (extracting the slope) the relationship between signal-to-noise ratio and dispersion to isolate meaningful

features. In the second stage, supervised machine learning is employed to classify the DPGs. Devine et al. analyzed 350 MHz drift scan data acquired by the Green Bank Telescope (GBT) in May-August 2011 (42 405 observations) and derived comprehensive classification for an unlabeled dataset of more than 1.5 million DPGs.

Machine learning methods have notably improved the efficiency of pulsar identification, but they require a substantial amount of labeled data, consuming time and computational resources. For this reason, scientists are now evaluating the potential of deep learning methods for pulsar identification. Balakrishnan et al. introduced the semisupervised generative adversarial network (SGAN) model (Fig. 2) that outperforms supervised machine learning methods relying primarily on unlabeled datasets [27] . Experimental results demonstrated that the model achieved high accuracy and average F1-score both equal to 94.9% when trained on 100 labeled and 5 000 unlabeled pulsar candidates. The final model version, trained on the "HTRU-S Lowlat Survey" dataset, achieved accuracy and average F1-score of 99.2% for both, with a recall rate of 99.7%. Subsequently, Yin et al. proposed a pulsar candidate identification framework that combines the Deep Convolutional Generative Adversarial Network (DCGAN) and deep aggregated residual network (ResNeXt) architectures [28] . To minimize sample imbalance, DCGAN 3) [29] . This model treats each diagnostic image of a pulsar candidate as a modality and successively applies multiple convolutional neural networks (CNNs) to extract meaningful features from the diagnostic images. After feature fusion, MFPIM [27] . [29] .

www.ati.ac.cn yields commonalities between different modalities in high-dimensional space, ensuring that the model has successfully extracted and used all complementary information contained in the diagnostic images. This approach demonstrably outperformed current supervised learning algorithms in terms of classification performance. Experi-mental results yielded a model accuracy higher than 98% for pulsar identification from the FAST dataset and accuracy and F1-score both exceeding 99% for the HTRU dataset.

Artificial Neural Networks (ANN)

Eatough et al. [30] 2010 A technique was proposed to automatically identify reliable pulsar candidates from pulsar surveys using artificial neural networks. Zhu et al. [31] 2014 A new artificial intelligence program, the Pulsar Image Classification System (PICS), has been developed to identify pulsars.

Support Vector Machine (SVM)

Devine et al. [26] 2016 A new two-stage method has been proposed to identify and classify dispersed pulse groups (DPGs) in the output of single pulse searches. Guo et al. [32] 2019 A framework that combines Deep Convolutional Generative Adversarial Networks (DCGAN) with Support Vector Machines (SVM) has been proposed.

Decision Tree (DT)

Lyon et al. [33] 2016 A novel online processing method based on decision trees has been proposed for the rapid identification of pulsar candidates. Tan et al. [34] 2017 An ensemble classifier consisting of five different decision trees has been developed to identify pulsar candidates in the LOTAAS dataset.

Convolutional Neural Networks (CNN)

Wang et al. [35] 2019 A deep convolutional neural network (CNN) with 11 layers has been designed for classifying pulsar candidates. Zeng et al. [36] 2020 A cascaded convolutional neural network (CCNN) has been designed for identifying FAST pulsar candidate data. Liu et al.

[29] 2023 A multi-modal fusion-based pulsar identification model (MFPIM) has been proposed to enhance the efficiency and accuracy of pulsar candidate recognition.

Deep Residual Networks (DRN)

Liu et al. [37] 2021 A 14-layer deep residual network (ResNet) was designed for the classification of pulsar candidates. Zhao et al. [38] 2022

The AdaBoost-MICNN framework was proposed to address the issue of poor performance in pulsar candidate identification algorithms caused by sample imbalance.

3.FRB SEARCH

FRBs are deep-space astrophysical phenomena characterized by short and intense emissions ("explosions") of electromagnetic radiation that last for milliseconds. Despite their extremely short duration, they exhibit exceptionally high brightness, with an energy release equivalent to the daily energy output of the Sun. The instantaneous radiation flux can reach tens of Jy [39]. Lorimer et al. first discovered FRBs while analyzing historical data from the Australian Parkes Telescope pulsar survey [40]. Thornton et al. defined the term "FRB" for such events, upon discovering four additional radio burst dispersion measurements in new survey data from the Parkes Telescope [41]. In April 2020, the first galactic (within the Milky Way) FRB was detected. Named FRB 20200428, it was generated by a magnetar (SGR J1935+2154) with an extremely strong magnetic field [42]. This unprecedented discovery supported the hypothesis that FRBs detected at cosmological distances were produced by magnetars. Since the first discovery, FRBs have been studied extensively. Standard FRB search methods involve visual identification from large astronomical observation datasets. Rapid advances in AI technology will enable real-time FRB search and multifrequency tracking observations. The random forest (RF) is a supervised machine learning algorithm. It can be described as an ensemble of decision trees that collectively form a robust classifier or regressor [43] . Wagstaff et al. improved the preexisting Very Long Baseline Array fast radio transients (V-FASTR) experiment by designing a machine learning classifier, derived from the RF algorithm, to automatically filter out known types of pulses [44] . By extracting specific data features and applying the RF model, candidate objects are classified into predefined RFI (Section 6) and FRB categories. The model achieved a classification accuracy of 98.6% for historical data and 99% or better for new observational data.

Single-pulse observation sensitivity is limited by RFI. To minimize this problem, Michilli et al. introduced the single-pulse searcher (SPS), an RF-derived machine learning classifier [45] .

Its purpose is to identify astrophysical signals in strong RFI environments. After optimization, the SPS was able to process observational data from the Low Frequency Array (LOFAR) Tied-Array All-Sky Survey (LOTAAS) and to accurately identify pulses from FRB source FRB 121102. Furthermore, Farah et al. developed an RF-based candidate classification pipeline for the Molonglo Radio Telescope using low-latency machine learning (Fig. 4) [46] . With this pipeline, they detected six FRBs from an accumulated dataset representing 344 search days between June 2017 and December 2018.

CNNs are among the feedforward neural networks most widely used for deep learning studies, with major achievements in image processing. A CNN notably reduces the number of parameters that need to be trained by optimizing the use of shared parameters and local connectivity through convolution and pooling operations. This, in turn, minimizes computational and storage costs [47] . Recently, deep learning techniques relying on CNNs have increasingly been applied to FRB searches, effectively solving the low-efficiency problem of standard visual screening for FRB candidates.

Cabrera-Vives et al. first applied a CNN to detect astronomical transient events [48]. Their model was designed to improve over the Deep High Cadence Transient Survey (Deep-HiTS) detection framework, itself implementing rotation-invariant deep CNNs was used to identify real sources in transient candidate images with high temporal resolution from the HiTS dataset. The model was validated and demonstrated higher performance (accuracy of 99.45% ± 0.03%) than simple RF models, with errors reduced by approximately half. Connor and van Leeuwen also applied deep learning to the single-pulse classification problem and developed a deep neural network (DNN) within a hierarchical framework (Fig. 5) [49] , to sort events by their probability of being astrophysical transients. Input to the DNN of Connor and van Leeuwen can include one or more data products (e.g., dynamic spectra and multibeam information) on which the neural network is simultaneously trained. FRB simulations provided the model with a larger and more diverse training set than when using single pulses from pulsars as training samples. Agarwal et al. proposed 11 deep learning models to classify FRB and RFI candidates [50] . Using transfer-learning techniques, they trained the stateof-the-art Fast Extragalactic Transient Candidate Hunter (FETCH) model on frequency-time and dispersion measure-time images using data from L-band observations of the GBT and from the 20 m telescope also located at the Green Bank Observatory. All models achieved accuracies and recall rates higher than 99.5%.

4.GW DETECTION

GW is currently a major research topic in physics and astronomy. It represents spacetime ripples generated by the acceleration of massive objects; its frequencies extend over an extremely broad range (nanohertz-kilohertz) that covers numerous physical processes and astronomical phenomena. Detecting GW with frequencies as low as the nanohertz is important for astronomers to understand cosmic structure and growth, evolution, and merging processes of large celestial bodies, such as supermassive black holes. It also provides physicists with information on fundamental principles of spacetime [51] . However, accurately detecting and analyzing GW signals from complex background noise is a challenging task. Application of AI technology has been successful in distinguishing genuine GW signals from false signals caused by other noise sources.

An artificial neural network (ANN) is a nonlinear, adaptive, information-processing model consisting of numerous interconnected processing units. It comprises an input layer, one or more hidden layers for data processing, and an output layer [52] . Data from external sources are collected in the input layer, processed in the hidden layers, and then the output layer provides one or more results depending on the function of the specific network. Kim et al. employed an ANN to search for GW signals related to short gamma-ray bursts [53]. They input multidimensional samples composed of statistical and physical quantities from a coherent search pipeline into their ANN to distinguish simulated GW signals from the background noise.

Experimental results showed that their ANN improved the efficiency of GW signal classification relatively to standard methods. Mytidis et al. studied the applicability to r-mode GW detection of three machine learning algorithms: ANNs, support vector machines (SVM), and constrained subspace classifiers [54]. They confirmed that machine learning algorithms achieve higher detection efficiency than standard methods. Krastev used deep learning techniques for rapid identification of transient GW signals from binary neutron star mergers (Fig. 6) [55] . The author proved that a deep CNN Fig. 4. Candidate classification pipeline using a low-latency machine learning algorithm [46]. trained on 100 000 data samples can swiftly identify such

signals and distinguish them from noise and binary black hole merger signals, thus validating the potential of ANNs for real-time GW detection. Mogushi et al. developed the NNETFIX ("A Neural NETwork to 'FIX' GW signals coincident with short-duration glitches in detector data") algorithm [56]. NNETFIX uses ANNs to denoise transient GW signals by reconstructing missing data portions caused by overlapping glitches, thereby enabling recalculation of the astrophysical signal localization. Experimental results demonstrated that NNETFIX successfully reconstructed most interferometer single-peak signals, with signal-to-noise ratio greater than 20, from binary black hole mergers. Wardega et al. proposed an ANN-based method for optical transient detection [57] . They created two ANN models: a CNN and a dense layer network, then tested them on samples generated from real images. Experimental results indicated an accuracy of 98% for this method. In the subsequent TOROS (Transient Robotic Observatory of the South) collaborative project following the gravity wave event GW 170104, the CNN and dense layer network achieved accuracies of 91.8% and 91%, respec- Fig. 6 . GW signals detection from binary neutron star mergers using artificial neural networks (ANN) [55] .

tively.

The advantage of deep learning over standard machine learning is its capacity to manage more complex data and tasks. Deep learning models can extract features automatically and execute classification, regression, and clustering tasks. Training the models on large datasets also improves their accuracy and generalization capabilities. In contrast, standard machine learning requires manual feature extraction and model training [58] . Therefore, deep learning methods have been applied to GW detection, with notable results. George and Huerta introduced an innovative end-to-end time-series signal processing method called "deep filtering", derived from two DNNs, with which weak time-series signals with non-Gaussian noise are rapidly detected and processed [59] . This method has been applied to detection and parameter estimation of binary black hole mergers. It expands the range of GW signals detection by ground-based instruments. Moreover, Wei and Huerta demonstrated that CNN-based deep learning algorithms are powerful enough to analyze GW despite their non-Gaussian and non-stationary nature [60] .

They illustrated the denoising capabilities of such algorithms by analyzing GW signals generated by binary black hole mergers GW150914, GW170104, GW170608, and GW170814. Chan et al. applied CNNs to GW signals from core-collapse supernovae (Fig. 7) [61] .

Additionally, by simulating time series of GW detector observations, they proved that CNNs relying on the explosion mechanism can detect and classify GW signals even weaker than the background noise. The true-positive rates for waveforms R3E1AC and R4E1FC l at 60 kiloparsecs (kpc) were 52% and 83%, respectively. For waveforms s20 and SFHx at 10 kpc, true-positive rates were 70% √ Hz √ Hz and 93%, respectively. Furthermore, Beheshtipour and Papa implemented a deep learning network model on the basis of a deep learning network model called "Mask R-CNN" to identify candidate signal clusters from continuous GW search output data, then evaluated model performance [62] . For high-frequency signals, this model achieved detection efficiencies higher than 97% with very low false-alarm rates. For low-frequency signals, the model maintained reasonable detection efficiency. On this basis, a second clustering network was designed to identify candidate signal clusters from weak signals [63] . Using the O 2 search data release from the Laser Interferometer Gravitational-Wave Observatory available from the "Einstein@Home" initiative (https://einsteinathome.org/), the cascading architecture of these two networks was validated to identify continuous GW candidate signal clusters with large differences in signal strengths.

The cascading network demonstrated excellent performance, achieving an accuracy of 92% at a normalized signal amplitude of 1/60 and exceeding 80% for weak signals at a normalized amplitude of 1/80.

5.SPECTRAL CLASSIFICATION

A fundamental task in stellar spectroscopy is systematic spectral classification. In particular, subtype classification is very important to study stellar evolution and identify rare celestial objects [64] . The considerably large astronomical spectral datasets contain information on unique celestial objects of major scientific significance, such as supernovae. Supernovae are violent explosions that occur in the late evolutionary stages of sufficiently massive stars; their extreme brightness exceeds that of their entire [61] .

host galaxy. Supernovae are among the most dramatic known astronomical events.

The study of supernovae is central to the field of astrophysics; their discovery represented a crucial step in the development of stellar evolution theories. Supernova research is also a cross-disciplinary science focusing on an object at the intersection of astrophysics, as a major stage of stellar evolution, and physics, as a natural laboratory to study extreme environments [65] . With observatories operating continuously and their equipment regularly upgraded, the available number of recorded spectra is constantly increasing and represents a major challenge for manual processing. The introduction of AI technology should provide notable support to improve the efficiency of spectral classification. Machine-learning algorithms have been used in astrophysical spectroscopy since their early stages of development, with promising results. The backpropagation neural network (BPNN) is a widely used neural network model trained with an error backpropagation algorithm. Because of its reduced training time, it is among the most common neural network models [66]. For example, Gulati et al. and von Hippel et al. applied two-layer BPNNs to stellar spectral classification, achieving good results in the optical and ultraviolet spectral domains at medium and low resolutions [64,67] . However, neuron selection markedly influences classification accuracy. Bailer-Jones et al. employed multiple BPNNs to construct a classification model that correctly classified 95% or more of dwarf and giant stars by luminosity, with high statistical confidence [68] . Additionally, they demonstrated that PCA was highly effective to compress spectra, with a compression ratio exceeding 30 and minimal loss of information.

Thus, Fuentes and Gulati combined MultiLayer Perceptron (MLP) neural networks with a PCA algorithm for subtype classification of stellar spectra and achieved classification confidence levels exceeding 95% for dwarf and giant star luminosities [69] . They also applied PCA for feature extraction from stellar spectra, with crucial results for feature extraction methods. Bazarghan proposed an ANNbased self-organizing map algorithm that does not require training data to classify stellar spectra [70] . Tested on 158 stellar spectra from the Jacoby, Hunter, and Christian library, this algorithm achieved a classification accuracy of 92.4%. Furthermore, Navaro et al. introduced an ANN system trained on a set of line-strength indices derived from temperature and luminosity-sensitive spectra (Fig. 8) [71]. This system enabled the classification of low signal-to-noise spectra, but it can be applied to other spectra with similar or higher signal-to-noise ratios. Finally, Bu et al. applied isometric feature mapping, SVM, and locally linear embedding algorithms to classify stellar spectra [72,73] .

Additionally, Cai et al. implemented "weighted frequent pattern trees" to establish association rules between stellar spectra [74].

They applied the information entropy theory to determine single-attribute weights of stellar spectra, using a compromise between geometric mean and maximum values to calculate the weights for multilevel item sets. Experimental results indicated that this method successfully retrieved the main stellar spectral type characteristics. Liu et al. investigated large-scale spectral data classification with a nonlinear ensemble classification method, the "non-linearly assembling learning machine" [75] . In this method, a large-scale spectral dataset is initially divided Input: 18 indices Fig. 8. ANN system for spectrum classification [71] .

Traditional classifiers are applied to each subset to produce partial results, subsequently aggregated to yield the final classification. Machine learning methods have markedly improved data processing speed, but manual feature extraction has simultaneously increased analysis complexity. Therefore, a substantial interest in further enhancing spectral classification performance exists in the scientific community. Jankov and Prochaska proposed the CNN-based "spectral image typer" to classify stellar spectral data [76] . This algorithm achieved an accuracy of 98.7% on both the validation and image-testing datasets. Additionally, Bu et al. introduced a method combining CNNs and SVMs (CNN+SVM method) to classify hot subdwarfs within LAMOST spectra [77]. Applying this method to a sample from the LAMOST Data Release 4 dataset yielded an F1score of 76.98%, demonstrating that the CNN+SVM method is highly suitable to identify hot subdwarfs in large-scale spectral surveys. Moreover, Sharma et al. presented a CNN-based automated stellar spectral classification method (Fig. 9), reducing classification errors to 1.23 spectral subclasses [78] . The pre-trained model was further applied to the Sloan Digital Sky Survey (SDSS) database to classify stellar spectra with signal-to-noise ratios greater than 20. To solve the problem of simultaneous spectral and luminosity types classification, Hong et al. introduced the "Classification model of Stellar Spectral type and Luminosity type based on Convolution Neural Network" [79] . In this model, CNNs allow for spectral feature extraction and incorporate attention blocks to discriminate between features and focus on essential ones; pooling operations are then conducted for dimensionality reduction, thereby compressing model parameters; finally, the fully connected layers are used for feature learning and stellar classification. In experiments with the LAM-OST Data Release 5 dataset (more than 71 282 spectra, each with more than 3 000 features), this model achieved an accuracy of 92.04% for spectral type classification and 83.91% for simultaneous spectral and luminosity types classification.

Finally, a "transformer" is a deep learning model architecture initially proposed by Vaswani et al. in 2017 and primarily applied to language processing. Application of transformer networks to astronomical research has gradually increased, primarily for astronomical data management and related issues. Liu and Wang implemented a high-performance hybrid deep learning model, the Bidirectional Encoder-Decoder Transformer (BERT)-CNN, to evaluate transformer performance for spectral classification [79] . The BERT-CNN model first uses stellar spectral data as input, then extracts feature vectors from the spectra with the transformer module before feeding these feature vectors into the CNN module to obtain classification results with a classifier called "softmax". Experimental results demonstrated that the BERT-CNN model effectively improved spectral classification performance, providing a valuable tool for application of deep learning concepts to astronomical research.

6. RADIO FREQUENCY INTERFERENCE MITIGATION

RFI is a disturbance that affects a communication or 10 www.ati.ac.cn observation system, for example by conduction or energy emissions in the radio spectral range.

RFI is generated by external sources such as human daily activities or telecommunications. During radio astronomy observations, various forms of RFI occur, interfering with the target signal by entering radio telescope receivers, reducing the observation signal-to-noise ratio, and introducing errors into the recorded data. With continuous advances in science and technology, an increasing number of radio frequencies is allocated to daily activities.

The artificial radio frequency signals thus generated cause severe interferences to deep-space astronomical signals, increasing the electromagnetic background complexity of telescope environments [80] . RFI mitigation is a major issue in signal processing for radio astronomy and its mitigation is essential for sensitive astronomical measurements such as pulsar timing.

However, standard RFI mitigation methods are affected by low efficiency and signal loss.

AI technology can efficiently identify and suppress RFI, thereby enhancing the signal-to-noise ratio of astronomical observation data.

Akeret et al. established a deep learning method to eliminate RFI signals with a specific type of CNN, U-Net, in order to discriminate between astronomical signals and RFI features in two-dimensional time-series observation data from radio telescopes [81] .

The study of such time series was the first application of deep learning technology to RFI mitigation. Czech et al. proposed a method to classify transient (temporal domain) RFI sources using deep learning techniques, by combining the CNN and long short-term memory (LSTM) architectures (CNN+LSTM method) [82] . They tested their method on a dataset containing 63 130 transient signals from eight common RFI sources, achieving good classification accuracy.

Kerrigan et al. presented a deep fully convolutional neural network (DFCN) to predict RFI by combining interferometric amplitude and phase measurements of signal visibility (Fig. 10) [83] .

Their experiments demonstrated a markedly improved time-efficiency for RFI prediction, with a comparable RFI identification rate. Liang et al. designed an algorithm to detect and extract HI (neutral atomic hydrogen) galactic signals using the Mask R-CNN network in combination with the PointRend method. They created a realistic dataset for neural network training and testing by simulating galaxy signals and RFI as potentially observed by FAST and selected the best-performing neural network architecture from an initial set of five [84] . This architecture successfully segmented HI galactic signals from time-domain ordered data contaminated by RFI, achieving an accuracy of 98.64% and a recall rate of 93.59%. Sun et al. proposed a robust CNN model for RFI identification [85] . They overlaid artificial RFI on data simulated by SKA1-LOW, creating three visibility function datasets. Experimental results indicated a satisfactory Area Under the Curve (AUC) of 0.93, demonstrating high accuracy and precision. To further assess model effectiveness, they also conducted RFI analyses with real observational data from LOFAR and the Meer Karoo Array Telescope (MeerKAT). The results indicated excellent performance, with overall effectiveness comparable to that of the AOFlagger [86]

(RASCIL) with the RFI simulation method from the Hydrogen Epoch of Reionization Array (HERA) telescope to generate RFI-affected simulated SKA1-LOW observations [87] . The CNN was trained on a subset of simulated data and subsequently tested on three additional datasets subjected to RFI with different intensities and waveform.

Results indicated effective RFI recognition by the proposed model from the simulated SKA1-LOW data, with an AUC of 0.93.

7. CONCLUSION

This review described representative astronomical applications of AI technology: pulsar candidate identification, FRB and GW detection, stellar spectral classification, and RFI mitigation. AI technology allows scientists to rapidly and accurately extract meaningful information from large-size astronomical observation datasets, thereby enabling deeper studies of the universe and celestial structures. Benefiting from the rapid development of information technology, AI technology is increasingly applied to astronomical researches. However, despite its usefulness, its limitations must be considered; these include interpretability (difficulty in analyzing the physical meaning of the results) and uncertainty (limited confidence level).

Future work should focus on improving the interpretability and credibility of AI models by employing advanced methods to interpret extracted features and predictive outcomes, with the purpose of improving their usefulness for astronomical researches. Furthermore, dedicated training datasets should be established to improve the usability of AI models. The emergence of generative AI systems such as ChatGPT will likely transform scientific research considerably.

The means of controlling and optimizing these advanced technologies for useful application to astronomical research should be carefully assessed and planned.

With continuous hardware performance improvements and gradual refinement of deep learning algorithms, deep learning techniques and neural networks will become increasingly prominent in astronomical data processing. Scientists expect the next generation of network architectures and training methods to improve identification accuracy for complex data patterns. Moreover, hardware performance improvements will allow the processing of increasingly larger and more diverse astronomical datasets.

In terms of network architecture, future designs will produce specialized and efficient neural networks, more adapted to the specificities of astronomical datasets and possibly including innovative structures such as CNNs, recurrent neural networks, or attention mechanisms designed to better capture spatiotemporal relationships, sequence information, and astronomy-specific features. In training methods, the introduction of new optimization techniques and adaptive learning strategies will enable researchers to account more effectively for noise, variations, and com-plex structures present in astronomical data. Moreover, the development of new machine learning techniques, such as transfer learning or reinforcement learning, should provide better solutions for model training with limited volumes of labeled data, thereby improving model generalization performance. Additionally, application of deep learning to astronomical data processing should enable scientists to investigate the joint processing of multi-modal data, for example, multiwavelength or multisource datasets, and to extract more comprehensive information from the observations. This holistic use of multi-modal data adopts a promising perspective toward a deeper understanding of astronomical phenomena and increased synergy between datasets. Overall, the development prospects of AI technologies for astronomical data processing are high. They will provide astronomers with better tools for deeper study of the universe.

AUTHOR CONTRIBUTIONS

Hailong Zhang conceived the ideas. Jie Wang summarized the results of the survey.

Yazhou Zhang provided the content of the Section 2 and 3. Xu Du provided the content of the Section 4. Han Wu provided the content of the Section 5. Ting Zhang provided the content of the Section 6. All authors contributed substantially to revisions. 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.