

A Review of Three Types of Neural Networks: Convolutional, Recurrent, and Recursive

Zheng Tang*

*Department of Electrical Engineering and Automation, Hefei University of Technology, Xuancheng Campus,
Xuancheng 242000, China*

**Corresponding author: Zheng Tang.*

Abstract

Artificial neural networks are a core technology in the field of nonlinear system modeling and intelligentization. Convolutional, Recurrent, and Recurrent Neural Networks, as their three core branches, each have unique structural and performance characteristics, and related research urgently needs systematic review and integration. This article adopts the methods of literature combing and comparative analysis to comprehensively sort out the development vein of the three types of neural networks, analyze their structural characteristics and applicable scenarios, explore the technical pain points and optimization paths of various networks, and look forward to the future development direction. The innovation of this research lies in the research results of the development and application of the three types of neural networks in the system, clearly defining the technical characteristics, differences between advantages and disadvantages and integration application modes of each network, and clarifying their respective optimization directions. The study found that the three types of neural networks have their own advantages and disadvantages, complementary adaptation, and have irreplaceable application value in computer vision, natural language processing and other fields. In the future, its development needs to focus on optimizing its own technical defects, improving the ability of model generalization, further expand the application boundaries through deep integration with other network models and learning mechanisms, and promote the application of intelligent technology in complex scenarios.

Keywords

convolutional neural networks, recurrent neural networks, deep learning, network structure, feature extraction, temporal data, model fusion

1. Introduction

Big data and high-performance computing have driven the evolution of artificial neural networks toward deep architectures, making them a core technology in the field of nonlinear system modeling and intelligence. Convolutional, recurrent, and recurrent neural networks, as their three core branches, have demonstrated irreplaceable application value in many fields. At present, the academic community has formed rich research results, including the combing of the development vein and technical pain points of a single network by Zhou Feiyan, Yang Li and other scholars, and the exploration of network integration applications by Liang

Chao and others, Zhang Lu, etc., and innovative research on the combination of neural networks with other mechanisms. However, the existing research mostly focuses on the breakthrough of a single network, and lacks the system integration and comparative analysis of the three types of networks. Based on this, this article systematically sorts out the development process, structural characteristics and applicable scenarios of the three types of neural networks, compares their technical advantages and disadvantages, explores the development direction of optimization and integration, and provides theoretical reference for the research and development and implementation of intelligent technology.

2. The History of Three Types of Neural Networks

2.1 The History of Convolutional Neural Networks

The study of convolutional neural networks can be roughly divided into three stages, namely the proposal in the mid-20th century, the realization at the end of the 20th century, and the attention in the 21st century.

2.1.1 Proposed in the Mid-20th Century

In the early 1960s, Hubel and Wiesel put forward the concept of receptor field through the study of the visual cortex system of the cat's brain, and further discovered the hierarchical processing mechanism of information in the visual cortex pathway, thus winning the Nobel Prize in Physiology or Medicine [1]. From the late 1950s to the early 1960s, Rosenblatt added learning functions on the basis of the MP model, proposed a single-layer sensor model, and put the study of neural networks into practice for the first time. However, the single-layer sensor network model cannot deal with the linear indivisibility problem. Until 1986, Rumelhart and others proposed a multi-layer feedforward network trained by error reverse propagation algorithm - Back Propagation Network (BP network), which solved problems that some original single-layer sensors could not solve [2]. Then in 1980, Fukushima first proposed a theoretical model based on the sensing field Neo cognitron, which became the main learning method of early neural networks.

2.1.2 Realization at the End of the 20th Century

In 1998, LeNet-5 proposed by Lecun et al. adopts a gradient-based backpropagation algorithm for supervised training of the network. The trained network converts the original image into a series of feature diagrams through the alternately connected convolution layer and the downsampling layer, and finally classifies the feature expression of the image through the fully connected neural network. The convolutional kernel of the convolutional layer completes the function of sensing field, which can excite the local area information of the lower layer to a higher level through the convolutional nucleus [1].

2.1.3 Attention in the 21st Century

In 2006, Hinton and others published an article in Science. Their main views are: (1) The artificial neural network of multiple hidden layers has excellent characteristic learning ability; (2) it can effectively overcome the training difficulties of deep neural networks through "layer wise pre-training", which has led to the study of Deep Learning, and also set off another boom in artificial neural networks [1]. In 2012, Alex Net proposed by Krizhevsky and others won the championship in the image classification competition of the large image database ImageNet [19] with a huge advantage of exceeding the second place by 11% in accuracy, making convolutional neural networks the focus of the academic community [2]. In addition, convolutional neural networks continue to integrate with some traditional algorithms, coupled with the introduction of migration learning methods, which has rapidly expanded the application field of convolutional neural networks. Recently, the application of convolutional neural networks has further developed to a deeper level of artificial intelligence, such as natural language processing, speech recognition, etc., and has been widely used in the field of computer vision, such as image classification, target detection and target tracking.

2.2 History of Recurrent Neural Networks

Compared to convolutional neural networks (CNNs), recurrent neural networks (RNNs) have received less attention due to the difficulty in implementing their early algorithms. Jordan and Elman proposed the RNN framework in 1986 and 1990 respectively, called the Simple Recurrent Network (SRN), which is

considered the foundational version of the widely used RNNs today. Subsequent, more complex structures can be considered variations or extensions of SRN [3]. But nowadays, it is also combined with convolutional neural networks in some fields, thus achieving some new results. For example: the multi-point wind speed prediction method in the wind farm of the convolutional recurrent neural network [4] and the traffic flow prediction of the convolutional recurrent neural network of the space-time diagram [5].

2.3 The History of Recursive Neural Networks

Recursive neural networks were first born in the 1980s, and their ability to maintain past input memory has opened up new fields for the study of neural networks. Since the 1990s, scholars have been studying it. For example, Hochreiter and Schmidhuber proposed long-term and short-time memory networks, and Cho and others proposed an RNN improvement structure similar to LSTM: gated memory unit. Since the 21st century, with the massive data brought by the Internet + and the increasing improvement of computer performance, has begun to play an important role in more and more fields. Its joint application with convolutional neural networks has also achieved some results, such as traffic status index prediction based on the fusion of convolutional and recursive neural networks [6], natural language processing and computer vision [7].

3. An Introduction to the Structures of Three Types of Neural Networks

3.1 Introduction to the Structure of Convolutional Neural Networks

The convolutional neural network is composed of input layer, hidden layer and output layer. The hidden layer contains multiple repetitive structures composed of convolution alllayer and sub-sampling layer [8].

(1) The convolution layer (Conv), also known as the feature extraction layer, is mainly used to extract the features of images. It consists of a set of convolutional kernels, and the weight value of the convolutional kernel can be automatically learned and updated according to the target function. The convolutional layer is an important core of convolutional neural networks. (2) Pooling layer, also known as the downsampling layer, generally performs dimension reduction operations between two consecutive convolutional layers, which can effectively reduce the amount of model parameters and reduce the overfitting of the network. Generally, there are max pooling and average pooling, etc. (3) Nonlinear units are composed of nonlinear activation functions, which are divided into saturated nonlinear activation functions according to their characteristics, such as Sig mod function, tanh function, etc., and unsaturated nonlinear activation functions, such as ReLU function, Leaky ReLU function, etc. The nonlinear unit maps the output results of the convolutional layer, so that the neural network can arbitrarily approach any nonlinear function, which improves the characteristic expression ability of the model. (4) Batch normalization (BN) is to transform the input data into the standard normal distribution, so that the input value of the nonlinear unit falls into the value interval with a large gradient, avoid the problem of gradient disappearance, and accelerate the convergence speed of model training. (5) Fully connected layers (FC) are usually used for classification tasks at the tail of the neural network, which converts the two-dimensional vector output by the convolutional layer into one-dimensional vector. Because the full connection layer connects all the neuron weights in the upper layer, there are many redundant parameters and occupy a large number of hardware resources, which destroys the spatial structure of the image. At present, convolutional layers or global average pooling layers (GAP) are commonly used to replace the full connection layer. A typical CNN architecture usually alternates the convolutional layer with the pooled layer, and finally outputs the results through one or more fully connected layers. In some cases, the global flat pooled layer is used to replace the full connection layer, and operations such as batch normalization are added to further optimize the performance of CNN. As shown in Table 1, there are four typical models of convolutional neural networks, namely, the stacking structure model with only one trunk network, the network-in-network structure model that calculates and re-aggregates in parallel through multiple network branches at multiple scales, the residual structure model using the short-circuit mechanism network, and the attention mechanism model using channel attention and spatial attention [9].

Table 1: Comparative analysis of Four model structures

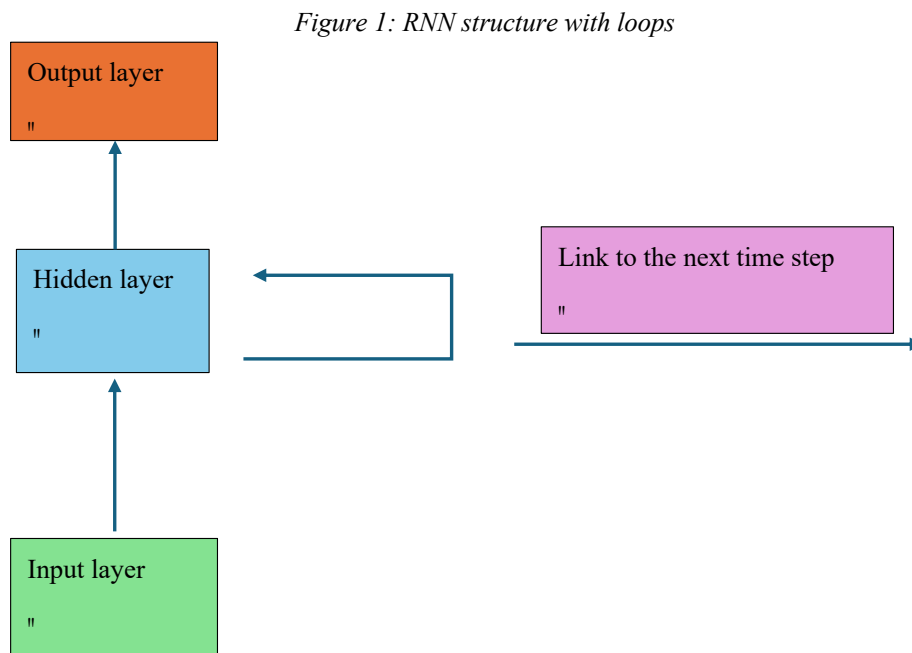
Model Name	Model Mechanism	Advantages and disadvantages	Application suggestions
------------	-----------------	------------------------------	-------------------------

Model Name	Model Mechanism	Advantages and disadvantages	Application suggestions
Stacked structure model	Only one backbone network	The network architecture is simple and easy to implement in both software and hardware, but the deeper layers are difficult to train.	Suitable for tasks involving smaller datasets
Net within a net structure model	Parallel computation and aggregation of multiple network branches across multiple scales	It can extract richer features and improve nonlinear representation capabilities, but it involves a large number of parameters and incurs high computational costs.	Avoid using it in systems with extremely limited resources whenever possible
Residual Structure model	Short-circuit mechanism network	It resolves the issues of deep network model construction and degradation, but exhibits the phenomenon of random depth.	Used for building deep neural networks
Attention Mechanism model	Channel Attention and Spatial Attention	This enhances the model's ability to autonomously extract features, but increases the model's complexity and computational load.	Use when you need to further enhance the performance and generalization ability of the original model

3.2 Introduction to the Structure of Recurrent Neural Network

Recurrent Neural Network (RNN) is mainly used to process sequence data. The hidden layers of RNN are connected, that is, the output of the current moment is not only related to the current input, but also depends on the output of the hidden layer at the previous moment [10]. Its biggest feature is that the output of neurons at a certain moment can be used as input to the neurons again. Compared with convolutional neural networks, this series network structure is very suitable for time series data and can maintain the dependency in the data. For expanded RNN, repeated structures can be obtained and the parameters in the network structure are shared, which greatly reduces the neural network parameters required for training [3].

Figure 1 shows the network structure of RNN. Through the loop connection on the hidden layer, the network state of the previous moment can be passed to the current moment, and the state of the current moment can also be passed to the next moment [3].



However, recurrent neural networks have the problems of gradient disappearance and gradient explosion. One solution is to change the stacking architecture between deep circular neural network layers to a two-way memory flow conversion mechanism to solve the shortcoming of memory state independence between layers.

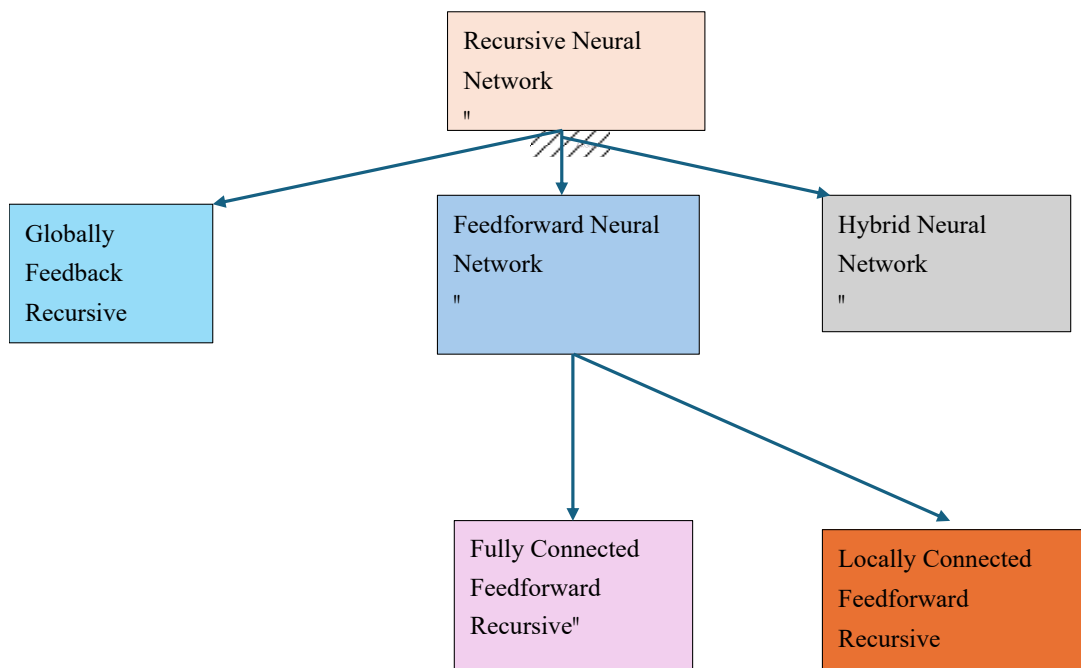
It provides an alternative fast path for the gradient flow from the output to the previous input of the long range, and introduces a self-attention mechanism inside the unit, so that the model can dynamically allocate importance, so that the hidden state of the current time step and the gradient highway unit output of the previous time step form a more effective relationship. This helps to improve the transmission and gradient flow of information and reduces the problem of gradient disappearance, so that the network can better capture complex patterns in the time series. However, under the action of the reverse propagation chain law, the gradient of the deeper the network decreases exponentially. In order to overcome this deep time dilemma, the gradient highway unit (GHU) can be introduced to the interlayer architecture, and the attention mechanism can be added to the unit, and the self-attention gradient highway unit (SAGHU) can be improved, which has brought a significant improvement to the performance of the model, effectively alleviated the gradient propagation problem in the deep LSTM structure, and alleviated the gradient disappearance problem [11].

3.3 Introduction to the Structure of Recursive Neural Networks

Although Recursive Neural Networks (RNNs) are not as well-suited for solving time-series data as Recurrent Neural Networks (RNNs), they excel in research areas involving continuous data, such as text, audio, and video, where model learning from sequential inputs is crucial. Therefore, RNNs are generally used in situations requiring sequential data processing. A typical characteristic of RNN architecture is recurrent connections, which allows the RNN to update its current state based on historical and current inputs, meaning all inputs influence future outputs [7]. Unfortunately, when the differences between related input data are large, full RNNs and selective RNNs cannot connect relevant information. To address “long-term dependencies,” Hochreiter et al. proposed Long Short-Term Memory (LSTM) to solve this problem [7]. Recursive networks are divided into locally connected and fully connected recursive neural networks. The essential difference between locally connected and fully connected recursive neural networks lies in the complexity of their network structures [12].

As shown in Figure 2, to put it simply, the types of recursive neural networks are divided into global feedback recursive networks, forward recursive networks, hybrid recursive networks, etc., while forward recursive networks are divided into full-connection forward recursive and local connection reverse recursion.

Figure 2: Types of Recursive Neural Networks



3.4 A Structural Comparison of Three Types of Artificial Neural Networks

As shown in Table 2, convolutional neural networks utilize convolutional kernels for feature extraction, downsampling to preserve information, sparse connections, and weight sharing. They involve large computational demands and are suitable for scenarios with large computational demands and fixed input image sizes, such as image processing.

Recurrent Neural Networks (RNNs), on the other hand, have interconnected hidden layers, so the output of a neuron at a certain moment can be used as input to the neuron again. They are suitable for time series data, can maintain the dependencies in the data, and can be used to solve time series data problems, or in conjunction with convolutional neural networks.

The output of a hidden layer node in a recursive neural network depends on the input of the current node and the value of the previous node, resulting in different input-output sequences and decreased accuracy due to long-term dependence. However, it is strong in terms of extracting time series data and generalization ability for continuous data research. It can be used to process time series data and applied to speech recognition, natural language understanding, etc.

Table 2: Comparison of Three Artificial Neural Network Models

Name	Mechanism	Structure characteristic	Function characteristics	Applicable Scenarios
Convolutional Neural Network	Convolutional kernel feature extraction, subsampling while preserving information	The computational load is large, and the input image size is fixed	Sparse connections, weight sharing	Scenarios involving image processing and other computationally intensive tasks where the input image dimensions are fixed
Recurrent Neural Network	The hidden layers of an RNN are interconnected, meaning that the output at the current moment is not only related to the current input but also depends on the output of the hidden layer at the previous moment	The output of a neuron at a given moment can be fed back into the neuron as input	Suitable for time series data; it preserves the dependencies within the data	Time series data, used in conjunction with convolutional neural networks
Recursive Neural Network	The output of a hidden layer node depends on the input to the current node and the value of the previous node	Different input and output sequences; accuracy decreases when processing long-term dependencies	Strong in areas such as the analysis of continuous data, feature extraction, and generalization ability	Used for processing time-series data, with applications in speech recognition, natural language understanding, and other fields

4. The Development Direction of Three Types of Neural Networks

4.1 The Development of Convolutional Neural Networks

Although the recognition rate of the traditional convolutional neural network model is already very high, by adding some specific noise points to the original picture, this artificial disturbance can easily make the neural network model misclassify the image. How to solve this problem and improve the generalization ability of the model is an urgent problem to be solved [9].

Convolutional neural networks can also be combined with other tools to improve the effect. For example, recently, Gao Xiaoping has proposed a deep learning model that combines Convolutional Neural Network (CNN), Long Short-Term Memory Network (LSTM) and attention mechanism (Attention), which is used for the accurate identification and classification of complex reservoir rock properties [13]. Through data cleaning, missing value supplementation and standardized processing, the quality of input data is improved.

Zhu Hongyu and others also proposed the CTC-Net model. The model enhances the characterization ability of mental fatigue characteristics and improves classification performance by collaboratively extracting the local time domain characteristics of the signal and the overall long-term dependence [14].

4.2 The Development of Recurrent Neural Networks

Time series data is a high-dimensional data type whose attributes change over time, reflecting the dynamic changes in user states (such as user interests) and movie states (such as movie popularity). For example, in daily life, users' preferences change with the maturation of their thinking or changes in their lifestyle, and their interest in specific types of programs may weaken or strengthen. Therefore, in addition to modeling the time dynamics, it is also necessary to overcome the hindsight bias in predicting ratings and to satisfy causal predictions. Gao Maoting et al. proposed a recommendation algorithm, R-RNN, which uses a variant of RNN, LSTM, to mine user states and movie states that change over time and generate predicted ratings. It has good robustness and relatively accurate rating predictions [13].

4.3 Development of Recursive Neural Networks

Fully connected recursive neural networks (RNNs) suffer from several drawbacks due to their large number of parameters: 1) network instability; 2) excessively long convergence time; and 3) training fluctuations [12]. Locally connected RNNs are better in this regard. On the other hand, the numerous parameters of fully connected RNNs allow them to store more historical data, enabling the identification and control of more complex nonlinear dynamic systems.

Recursive neural networks can also be used in combination with convolutional neural networks. For example, Zhang Lu, Liu Shuang, and others explored and verified that the model based on the fusion of convolutional and recursive neural networks performed better in long-term and congestion traffic state index prediction [6].

Wang Kang et al. designed a convolutional recursive multispectral image compression method based on spatial spectral feature extraction, which can effectively preserve and restore the spatial and spectral information of multispectral images [15].

5. Summary

This paper reviews three types of core neural networks: convolutional, recurrent, and recursive, sorts out their development context, structural features, applicable scenarios, and looks forward to their development directions. All three originated in the middle and late 20th century, gradually matured with the development of deep learning, and often integrated with other networks or mechanisms, playing a key role in the field of intelligence. Convolutional Neural Networks have developed through three stages: proposal, implementation, and attention. They extract spatial features by relying on convolutional kernels, pooling layers, etc. There are four typical models, which are the mainstream in computer vision. Recurrent Neural Networks achieve memory functions by connecting hidden layers and are adapted to time series data, but they have gradient problems, which can be alleviated by improving the architecture. Recursive Neural Networks process nested/continuous sequence data through recursive connections, divided into partial and fully connected types. LSTM has solved the long-term dependence problem and is suitable for fields such as speech recognition and natural language processing. The three types of networks each have their own advantages and disadvantages, complement each other and are suitable for different scenarios. Currently, they all have their own technical pain points. In the future, its development will all revolve around optimizing its own shortcomings and enhancing generalization capabilities. The core direction is to integrate with other models and mechanisms, expand application boundaries, and achieve performance breakthroughs in more practical scenarios.

References

- [1] Zhou, F., Jin, L., & Dong, J. (2017). A review of convolutional neural networks. *Chinese Journal of Computers*, 40(6), 1229–1251.

- [2] Li, Y., Hao, Z., & Lei, H. (2016). A review of convolutional neural networks. *Journal of Computer Applications*, 36(9), 2508–2515, 2565.
- [3] Yang, L., Wu, Y., Wang, J., & Zhou, C. (2018). A review of recurrent neural networks. *Journal of Computer Applications*, 38(S2), 1–6, 26.
- [4] Liang, C., Liu, Y., Zhou, J., & Li, J. (2021). Multi-point wind speed forecasting method in wind farms based on convolutional recurrent neural network. *Power System Technology*, 45(2). <https://doi.org/10.13335/j.1000-3673.pst.2020.0767>
- [5] Gu, Z., Chen, C., Zheng, J., & Wu, H. (2022). Traffic flow prediction based on spatio-temporal graph convolutional recurrent neural network. *Control and Decision*, 37(3). <https://doi.org/10.13195/j.kzyjc.2020.1445>
- [6] Zhang, L., Liu, S., & Tian, Y. (2021). Traffic state index prediction based on the fusion of convolutional and recurrent neural networks. *Traffic and Transportation*, 37(1).
- [7] Wang, Y., Liao, B., Peng, C., & Yang, X. (2021). A review of recurrent neural networks. *Journal of Jishou University (Natural Sciences Edition)*, 42(1). <https://doi.org/10.13438/j.cnki.jdzk.2021.01.007>
- [8] Liu, W., Liang, X., & Qu, H. (2016). Research on learning performance of convolutional neural networks with different pooling models. *Journal of Image and Graphics*, 21(9).
- [9] Yan, C., & Wang, C. (2021). Development and application of convolutional neural network models. *Journal of Frontiers of Computer Science and Technology*, 15(1).
- [10] Gao, M., & Xu, B. (2019). Recommendation algorithm based on recurrent neural network. *Computer Engineering*, 45(8). <https://doi.org/10.19678/j.issn.1000-3428.0051294>
- [11] Wang, Y. (2025). Research on neural network models for forest fire risk prediction [Master's thesis, Tianjin University of Technology]. <https://doi.org/10.27360/d.cnki.gtlgy.2024.000605>
- [12] Cong, S., & Dai, Y. (2004). Research on the structure of recurrent neural networks. *Journal of Computer Applications*, (8), 18–20, 27.
- [13] Gao, X., Zhu, H., Zhao, K., & Wang, J. (2026). Lithology identification method based on convolutional neural network-long short-term memory artificial neural network-self-attention mechanism. *Science Technology and Engineering*, 26(2).
- [14] Zhu, H., Zhen, Q., Chen, Q., & Wang, H. (2026). A novel mental fatigue detection method based on the fusion of single-channel EEG convolutional neural network and transformer. *Journal of Biomedical Engineering*.
- [15] Wang, K. (2025). Research on multispectral image compression algorithm based on convolutional recurrent neural network [Master's thesis, Nanjing University of Aeronautics and Astronautics]. <https://doi.org/10.27239/d.cnki.gnhhu.2023.003084>

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).