

Retraction

Retracted: Multimedia Data Processing Technology and Application Based on Deep Learning

Advances in Multimedia

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Advances in Multimedia. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] G. Li and W. Liu, "Multimedia Data Processing Technology and Application Based on Deep Learning," *Advances in Multimedia*, vol. 2023, Article ID 4184425, 15 pages, 2023.

Research Article

Multimedia Data Processing Technology and Application Based on Deep Learning

Guo Li  and Wei Liu

College of Mechanical and Electrical Engineering, Nanyang Normal University, Nanyang, Henan 473000, China

Correspondence should be addressed to Guo Li; lg_2018@nynu.edu.cn

Received 7 September 2022; Revised 28 September 2022; Accepted 22 March 2023; Published 7 April 2023

Academic Editor: Tao Zhou

Copyright © 2023 Guo Li and Wei Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With deep learning being widely used in various research fields, it is introduced into the research and analysis of multimedia data processing technology and application. First, the flow of multimedia data processing, the development of multimedia data, and the realization of multimedia data processing technology are explained and analyzed. Then, the related network results of deep learning (convolution network structure and countermeasure neural network structure) are put forward, and the image comparison of the activation function and the loss function of deep learning is analyzed, which provides functional algorithm support for the experimental analysis of deep learning in multimedia data processing technology. Finally, through the analysis of experimental data, it is concluded that deep learning has stronger advantages in the application research of multimedia data processing technology compared with other learning methods. In the multimedia data processing, the multimedia data processing technology is obviously superior to the data mining technology and data compression technology. Finally, under the support of deep learning data, we conclude that multimedia data processing technology is widely used and quoted in various fields. Therefore, with the development of multimedia, the amount of multimedia data is increasing; so, we should vigorously develop multimedia data processing technology in an all-round way.

1. Introduction

Deep network of deep learning has been successfully applied to unsupervised feature learning of single mode multimedia. In deep learning, this paper proposes a new application of deep network learning features in multiple modes and a series of tasks for multimodal learning and shows how to train deep network learning features to solve these tasks. Among them, how to learn the shared representation between patterns in deep learning and evaluate it on a unique task is shown [1]. In the training of deep learning, the mainstream method advocates the use of the random gradient descent method. Although the random gradient descent method is easy to implement, it is difficult to adjust and make it parallel. These problems make it challenging to develop, debug, and extend deep learning algorithms. In addition, it is proved in experiments that more complex off-the-shelf optimization methods, such as finite notation and conjugate gradient with line search, can significantly

simplify and speed up the process of the deep learning algorithm [2]. The deep learning method and its application in various signal and information processing tasks are widely used. The successful application of deep learning technology has been transformed into application fields. In-depth research can have an important impact on the application field and benefit from the recent research work, including multiscale in-depth research on natural language and word processing, information retrieval, and computer science [3]. Deep learning can use input structure to extract data-related features. Its purpose is to make these features more abstract, make their individual features more invariant to most of the changes that usually exist in the training distribution, and at the same time collectively preserve the information in the input as much as possible. And under ideal circumstances, the related algorithms of deep learning can train unknown mutation factors based on distribution [4]. Convolution neural network in deep learning has the convolution layer, pooling layer, and transposed convolution layer. It mainly

introduces after multilayer processing. The data features of research samples can be transformed from low features to high features. The properties of deep learning layer and the relationship between the convolution layer and the transposition layer of upgrade network are introduced in detail [5]. Deep learning studies the application of the new physical layer, mainly using network as automatic encoder. A new deep learning method is proposed, which takes network design as an end-to-end reconfiguration function, aiming at optimizing the components of transmitter and receiver through a method. Finally, the application of deep learning convolution neural network in modulation classification of original samples is demonstrated. Compared with the traditional modulation classification method relying on expert features, deep learning obtains competitive accuracy [6]. Multimedia data processing is accomplished by automatically assembling filter diagrams including filters that operate to perform processing functions on the data stream. Filter diagrams can be assembled by selecting the appropriate filters that can handle the data processing requirements of the desired data stream. Where the graph is constructed by selecting a set of filters, the filter comprises a suitable file reader compatible with the media type of the data stream for separating the multiplexed data, a decoder for decoding the encoded data, and a render for displaying or playing the data and combines these filters within the architecture of the filter graph to efficiently process the multimedia data [7]. With the increasingly complex and intelligent demand of multimedia services, the computing demand of multimedia data processing is also growing steadily. New multicore hardware architectures provide the required resources but writing parallel distributed applications is still a labor-intensive task compared to their sequentially opposite parts. Therefore, a multimedia data processing framework is proposed. An inherent limitation in the design of these processing frameworks is that they cannot represent any complex workload. Moreover, the dependency graph of the multimedia data processing framework is often limited to directed acyclic graph, even in a predetermined stage [8]. Correct analysis of multimedia data requires effective extraction and segmentation techniques. Among many computational intelligence methods, deep learning is most suitable for several tools and technologies containing intelligent concepts and principles and dedicated to object extraction, image segmentation, and edge detection using deep learning techniques with a wide range of real-life applications of images and multimedia data. Then, it introduces the basic knowledge of brain structure and learning and then introduces the key deep learning technologies, including evolutionary computation, neural network, fuzzy set and fuzzy logic, and rough set [9]. The multimedia network information processing system includes a main switch sending data from a server to a terminal and sends multimedia network data and multimedia network data protocol to a main processor. The large network processor and the multimedia network component group the data packets and send the data packets to the unidirectional optical fiber network. The data processing module receives the data broadcast by one-way optical fiber network, decrypts the data, and realizes data

exchange according to the multimedia network protocol [10]. The multimedia information system is a real-time transparent audio stream video format. The operating environment according to the embodiment does not require the participation of a video and audio stream generator or a client application. With this seamless solution, video and audio streams can be processed seamlessly and completely independently, regardless of the user's choice of a particular client application. In multimedia data processing embodiment, external services are primarily used to track new processing and add code to the processing [11]. The multimedia data processing mainly includes an input device for receiving input information containing one or more kinds of data, a data analyzer that checks each type of data constituting the input information and extracts data from the input information when the checked data is judged to be predetermined conversion target data to be subjected to data conversion processing, and a controller that performs conversion processing on the amount of the extracted data in accordance with a predetermined specification of the data type [12]. The efficient exchange of multimedia data in a data processing system establishes a sequential data stream comprising a plurality of variable length continuous segments, each of which comprises a plurality of data samples. Each data sample preferably includes a data set and a control structure or header for specifying a manner of interpreting the data set. The control structure preferably includes information on the size of the data collection, the resolution and duration of the data collection, the data collection method used, and the coding technique used [13]. The multimedia data processing device is mainly a template storage device, which stores multimedia template data. The multimedia template data comprise media data and scene data. The scene data define an output aspect of the media data, a data input means for inputting media data for transmission to be transmitted to a client terminal, a template designating means for designating multimedia template data for creating multimedia data using input media data for transmission from among the stored multimedia template data, and multimedia data creating means [14] for creating multimedia data by exchanging media data for exchange with input media data for transmission. Multimedia data processing is a key technology in the application fields of multimedia, network communication, and deep learning. Massive image, video, audio, and other multimedia data make it necessary to store and transmit multimedia data in compressed format. The application of processing technology makes it more difficult to directly process multimedia data. Therefore, the compressed domain processing of multimedia data has become an interesting research field [15].

2. Multimedia Data Processing Technology and Related Applications

2.1. Multimedia Data Processing-Related Processes. Multimedia processing technology adopts the client-serversystem-related mode. The internal process of multimedia application or multimedia information processing

technology requires the server to switch to the service mode of the server according to the request of multimedia information processing technology and provides feedback from the request processing and other aspects, i.e., receives the request, process, and return an error message, and finish processing the report. Finally, the request terminates the conversion process, destroys the object conversion, and releases the system resources. When performing the above steps, a control request can be sent to operate the communication process. For media data with time sequence, current play position, play cycle time, and other features can be set for startup, pause, recovery, and other related settings. The nontime series output may be required for editing and analysis depending on the consumption of still image data. The principle and function of multimedia information processing technology between multimedia application software and its function is to provide embedded data processing equipment of operating system management equipment, hardware and software of multimedia data, and management of other information resources. There are many kinds of multimedia data, which mainly include audio data, image and video data, and related text data. These different multimedia messages include all kinds of information formats that are popular. Multimedia data must be consistent with the data processing format. Multimedia application service mainly applies multimedia information technology, which ensures that client and application server request service by calling the interface of multimedia information technology to provide resource information; the other is to process the request and return the result to the application. Multimedia information processing technology supports a variety of popular data formats, including all multimedia formats in the list. The flow of multimedia data processing technology is shown in Figure 1. From the analysis of Figure 1, it can be seen that, first, the processing technology will create and set the relevant processing sequence and then carry out conversion preprocessing and start the conversion function of multimedia data processing technology. Among them, the conversion function of multimedia data processing technology is divided into four states, and different data types will carry out different processing modes. Next, the data are stopped from converting, and the corresponding result of multimedia data processing is obtained. Finally, the object data processed and converted are deleted by itself, so as to avoid data confusion and occupying the storage space of multimedia data processing.

2.2. Overview of Multimedia Data Processing Development. Multimedia data information is mainly composed of unstructured data information such as audio data information and video image data information and is growing rapidly. In the recent years, digital library, Internet, and other information resources have developed rapidly and become important information resources in people's daily life. The development of multimedia data processing technology is very difficult up to now, and the development of multimedia has relative limitations in the 20th century. From the early

1990s to the present, the processing mode has changed from multimedia processing to metadata-based multimedia content in the recent years. Multimedia processing combining semantics and content improved until the end of 1990s. The multimedia data processing technology in the 20th century is used to process a single media object, and the multimedia data processing technology in the 21st century is used to expand the processing of multiple media types. From the development trend of multimedia technology, we can see that it is from the supporting simple multimedia to processing complex media. Video, audio, images, animation, and other multimedia resources make the whole world of data resources more beautiful and no longer constrained and monotonous data information. With the rapid development of modern society, the amount of multimedia data information is also increasing. Facing the vast ocean of multimedia information and data, how to quickly and accurately obtain the information needed by the users from massive multimedia information resources is a very important research topic, and it is also the focus of database and multimedia research. The essence of multimedia data processing technology includes three aspects: multimedia information processing, high-dimensional indexing technology, and parallel query technology. The limitation of multimedia data processing technology is usually based on the underlying visual and audio objects used to describe images or other multimedia information. People are used to measuring their important achievements on the semantic level. Existing computer vision technology is difficult to adapt to the high-order semantic concepts of media content objects; so, content-based processing methods have poor accuracy. Because the concept of multimedia device embodies high-level semantics, the processing effect of content-based multimedia technology is not ideal, and it is difficult to apply in practice. In order to solve the problem of multimedia data response, a feedback mechanism is proposed to improve the processing efficiency of the media data content by changing the interrelated data between media. The development of the multimedia processing mode is shown in Figure 2. By analyzing Figure 2, we can know that multimedia has undergone four major changes. In the 20th century, multimedia data processing technology is mainly embodied as a single multimedia data type. And the multimedia metadata processing technology from 1970s to 1980s has changed into multimedia data content processing in the mid-1990s. Then, after the 20th century, multimedia data processing is embodied as a variety of media data types, mainly multimedia data semantic processing technology and antimedia processing technology.

2.3. Implementation Technology of Multimedia Data Processing. Multimedia data processing technology is a relative technology for multimedia complex information. Multimedia data processing technology is mainly through the image, video, audio, and other related information processing. By opening the image file to scan the multimedia image processing, the obtained information will appear on the user's screen in the form of an image. Multimedia video

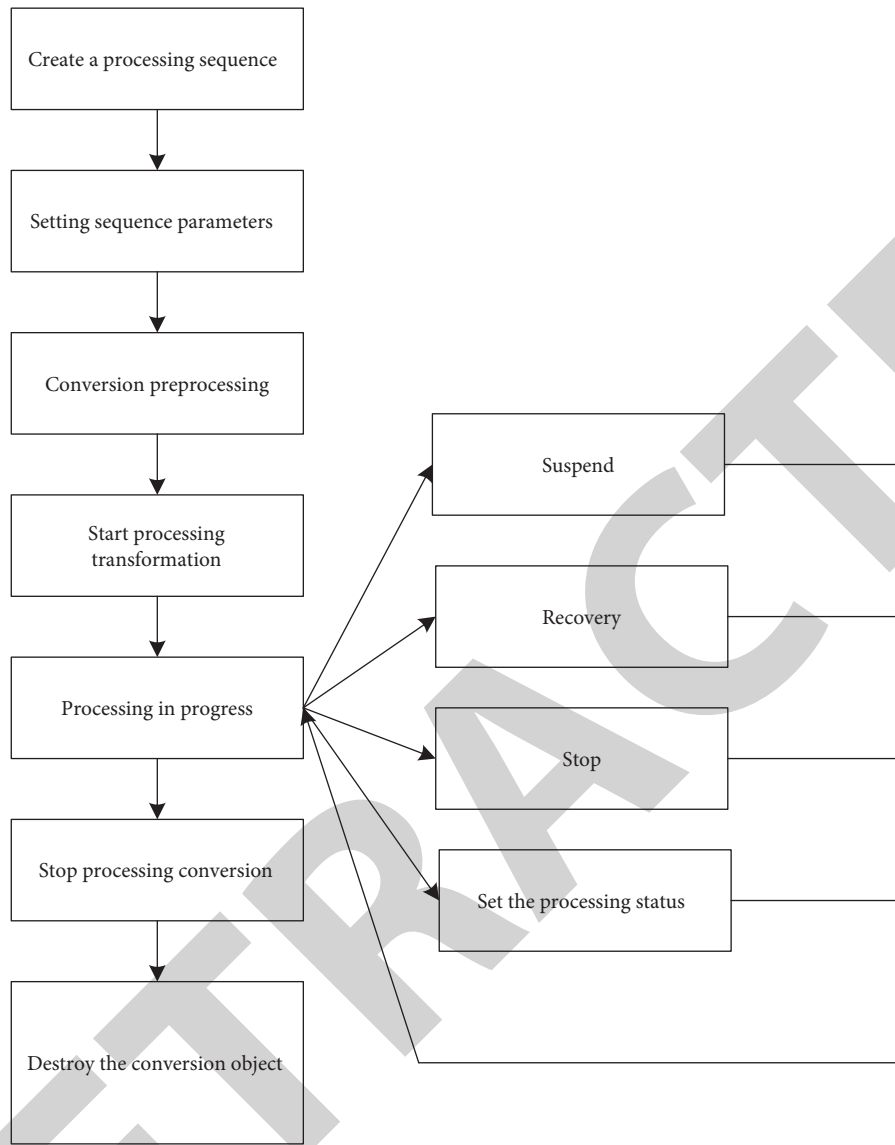


FIGURE 1: Flow of multimedia data processing technology.

processing searches for relevant information through a multimedia database, returns it to a temporary video file through a byte stream, and then reproduces the processing information in the file. Multimedia data processing is to store data in the multilevel storage of large-capacity text data, which can completely replace memory data types, then restore the processing method of the compressed control and restore and detect the original file information on the screen. The amount of multimedia data is very large and the storage methods of multimedia data are various. Currently, there are two common ways to store unstructured data directly outside of a database file. If you want to access and process multimedia data, you need to view the data and process it according to the site identifier. For example, image and audio data can be stored in objects in the form of file names. When the image needs to be displayed or transmitted, it is processed according to the name and process of the file. Unstructured data stored in applications can support

applications. Some multimedia database systems can support unstructured data storage, such as database systems. However, in the process of implementation, more programming techniques are needed to store data in unstructured databases. Once these problems are solved, more programming methods are needed. However, the advantage of users is openness. Users only know this multimedia data function and do not know how to store it, so they get different methods. The disadvantage of multimedia database is that programmers have complex program structure when storing data or retrieving and reading data, which requires a lot of execution time of processor. Multimedia data display certain information, such as images, images, and sounds. From the microscopic point of view, multimedia data are binary bytes. Byte data are combined with specific coding format, and its expression is realized by pattern recognition. Therefore, we have a way to manipulate multimedia files and manipulate data byte streams to save them or delete them

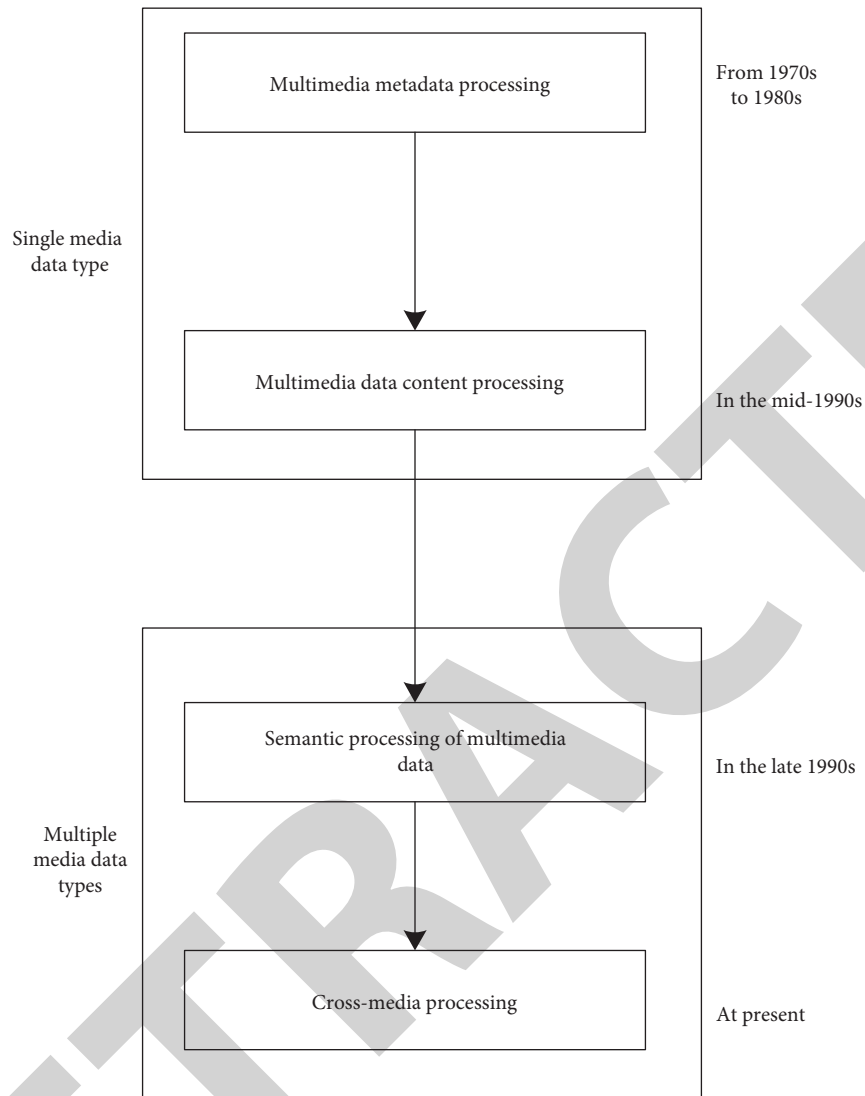


FIGURE 2: Development of the multimedia data processing mode.

from the database. The multimedia data processing access attribute process is shown in Figure 3. As can be seen from Figure 3, multimedia will first collect data from the multimedia database by obtaining the byte stream of data and then convert the byte stream into different data, thus displaying the data in the form of images, videos, audio, and so on. At the same time, multimedia can also express data byte stream through media expression software. Finally, the media will restore the byte stream in the database and transmit it to the multimedia attribute library so that the original byte stream can be used next time.

3. Related Algorithms of Deep Learning

3.1. Deep Learning Network Structure. Deep learning has been one of the topics that people care about since it was put forward. Deep learning neural network has been widely used in all aspects of social production, such as network model applied to large-scale image data processing, circular structure suitable for processing audio, voice, text and other

data, and image modeling, and other generation countermeasures networks that need to generate new samples. Therefore, we will focus on the volume neural network and generation countermeasure network commonly used in multimedia data processing. The full connection layer of neural network is generally used to integrate the information of the previous layer in neural network. The purpose of the convolution layer is mainly to extract the low-dimensional and high-dimensional features of the input image, in which the mathematical operations are mainly convolution operations, flat convolution, grouping convolution, and so on. The deep learning neural network is shown in Figure 4.

The complete whitening of the input in each layer of deep learning neural network requires a lot of computing resources and can be differentiated everywhere. Therefore, the batch normalization layer mainly optimizes the output and input characteristics of neural network structure. The calculation formula of deep learning batch normalization is as follows: In the deep learning neural network, normalization calculation can accelerate the convergence speed of the deep

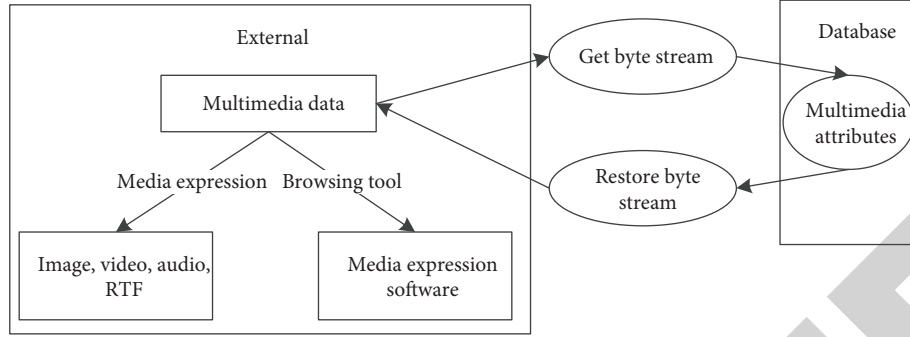


FIGURE 3: Access attribute process of multimedia data processing.

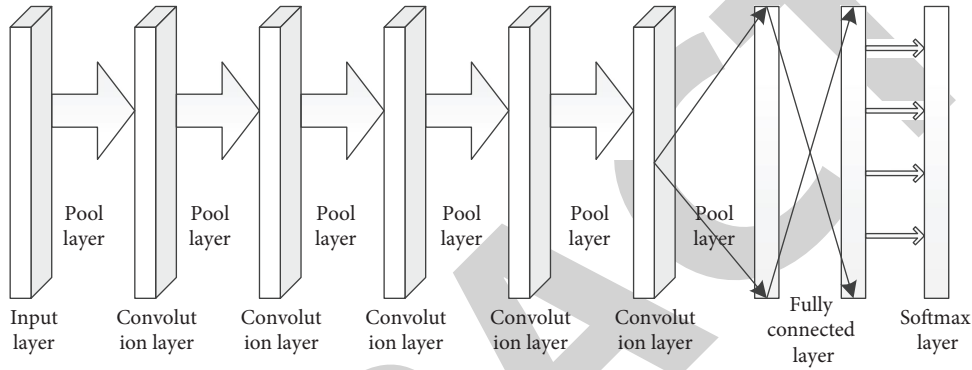


FIGURE 4: Related composition of deep learning convolution network.

learning network, and the required data type can be obtained by reducing and enlarging the normalization calculation of data, and then the content conversion of the deep learning neural network structure can be carried out. The normalized training calculation can store and calculate the output and input features of each scalar accurately in the convolution layer of neural network, which is beneficial to the accuracy of scalar features in the convolution layer and the independence of scalar data in neural network.

$$\bar{x} = \frac{x^{(K)} - E[x^{(K)}]}{\sqrt{\text{Var}[x^{(K)}]}}. \quad (1)$$

In deep learning neural network, expectation and variance are calculated on the training set. This normalization method accelerates the convergence speed of the network, but this simple normalization of each input in the input layer may change the content that the original input layer can represent; so, the batch normalization layer needs to be sure that the transformation added to the model can represent the content transformation. The calculation formula of deep learning normalized value scaling shift is as follows:

$$y^{(K)} = y^{(K)} * \bar{x}^{(K)} + \beta^{(K)}. \quad (2)$$

The learning of additional parameters in deep learning is synchronized with the learning of parameters in the original network model. The main purpose of introducing these additional parameters is to restore the representation ability lost by batch normalization operation of the neural network

model. Therefore, the calculation formula of deep number activation of deep learning is as follows:

$$\begin{aligned} y^{(K)} &= \sqrt{\text{Var}[x^{(K)}]}, \\ \beta^{(K)} &= E[x^{(K)}]. \end{aligned} \quad (3)$$

The second aspect is that the random gradient training can be carried out in batches at the same time, which improves the training speed of the network, and each batch will estimate the mean and variance of a single activation, respectively. Based on this, the parameters after batch normalization can still carry out gradient back propagation, and batch normalization is calculated based on the variance of each dimension of input data instead of the traditional joint variance calculation. Deep learning batch normalization transformation can be described as follows:

$$\{x_1, \dots, m\} N_{\gamma, \beta}: x \longrightarrow y. \quad (4)$$

Deep learning normalization transformation exists inside the transformation and the subnetwork input is formed, and then other processing is carried out by the original deep learning network model. The batch normalization layer can be introduced into the model, and parameters need to be learned in the transformation. In the convergence process, the gradient of batch normalized back propagation loss is also calculated, and the gradient of batch normalized parameters is calculated, which is calculated by using the chain rule. The deep learning normalized chain calculation formula is as follows: by

analyzing the chain calculation formula of deep learning, it can be concluded that the training speed of a single be improved by adopting the batch normalized layer chain calculation. Therefore, on the whole, using the chain calculation batch normalization layer can still improve the calculation speed of neural network. And the essence of the batch normalization layer in chain calculation is to normalize the input data of each layer of neural network. Therefore, the small changes of model parameters of deep learning neural network are prevented from being amplified, resulting in large deviation, and chain calculation can prevent the convergence speed of deep learning network and gradient explosion and disappearance.

$$\begin{aligned} \frac{\partial e}{\partial \alpha_B^2} &= \sum_{i=1}^m \frac{\partial e}{\partial x_i} \cdot (x_i - u_B) \cdot -\frac{1}{2}(\alpha_B^2 + \varepsilon)^{-3/2}, \\ \frac{\partial e}{\partial u_B} &= \sum_{i=1}^m \frac{\partial e}{\partial x_i} \cdot \left(\frac{-1}{\sqrt{\alpha_B^2 + \varepsilon}} \right) + \frac{\partial e}{\partial \alpha_B^2} \cdot \frac{\sum_{i=1}^m -2(x_i - u_B)}{m}, \\ \frac{\partial e}{\partial x_i} &= \frac{\partial e}{\partial x_i} \cdot \frac{1}{\sqrt{\alpha_B^2 + \varepsilon}} + \frac{\partial e}{\partial \alpha_B^2} \cdot \frac{\partial e}{\partial \alpha_B^2} \cdot \frac{2(x_i - u_B)}{m} + \frac{\partial e}{\partial u_B} \cdot \frac{1}{m} \\ \frac{\partial e}{\partial B} &= \sum_{i=1}^m \frac{\partial e}{\partial y_i}. \end{aligned} \quad (5)$$

The generated countermeasure network in deep learning needs to train the generated model at the same time, in which false data are generated mainly according to the data distribution of the training set, which is mainly used to judge whether the data come from the original sample data or other models. The training mode of the deep learning neural network model still adopts the back propagation mode. The GAN network model structure diagram of deep learning is shown in Figure 5. By analyzing the structure of the countermeasure network model in Figure 5, when random multimedia data enter the deep learning confrontation network, its countermeasure network generates a model G . Then, along with the real data stored in the countermeasure network model, it is included in the discrimination model D of the deep learning network structure. The model G will judge the authenticity of the multimedia data according to the relevant calculation structure and then return it to the model G of the countermeasure network. The discrimination will not stop until the multimedia data enter the neural network model for processing.

In the problem of deep learning model estimation, discriminant model learning determines the source of samples, generated model learning produces wrong samples and deceives discriminant model, and the discriminant model judges the source of samples. In the process of training and learning, both the generated model and the discriminant model are optimized until false data and real data cannot be distinguished. The sample discrimination calculation formula is as follows:

$$\begin{aligned} \min_G \max_D V(D, G) &= E_{x \sim P_{data}(x)} [\log D(x)] \\ &+ E_{z \sim P_z(z)} [\log(1 - D(G(z)))]. \end{aligned} \quad (6)$$

In the deep learning countermeasure network, when G and D are not limited in parameters, they can be trained by iterative or numerical methods. However, it is very difficult to optimize D in the convergence process. When the data set is small, it is easy to produce overfitting phenomenon. Therefore, the deep learning countermeasure network alternates between the step of optimizing D and the step of optimizing G . The specific calculation formula is as follows:

$$\nabla \theta_d \frac{1}{m} \sum_{i=1}^m [\log(1 - D(G(z^{(i)})))] \quad (7)$$

The main disadvantage of deep learning to generate confrontation network is that it is not clearly expressed, and in the training process, the generation model and the discrimination model will be trained simultaneously. Generating countermeasure network has the same advantages in calculation and fitting some data with sharp or even degraded distribution. Without considering the number of parameters and the calculation speed, the network model can easily converge to. In order to obtain the global optimal of the calculation model, the optimal discriminator given to the generator needs to be considered first, and the calculation formula of the deep learning optimal discriminator is as follows:

$$D_G^* = \frac{P_{data}(x)}{P_{data}(x) + p_g(x)}. \quad (8)$$

3.2. Deep Learning Correlation Function Operation. When the deep learning function is applied to the actual scene, the loss function only uses the features to extract the features of multimedia images without classification. Because the database of deep learning is dynamic, in this experiment, the images in the test set and the training data set do not overlap. In order to enhance the feature extraction ability of the model and the robustness of the test, the classification loss function is used to make the model have good discrimination ability, and the measurement loss function is used to enhance the feature expression and feature extraction ability of the model. In the application of deep learning recognition, after positioning, segmentation, and normalization, the multimedia is input into the feature extraction network and the output of the feature extraction network is taken as the vector representation of the multimedia. By measuring the Euclidean distance between the currently input multimedia and the registered Euclidean distance in the database, it can be judged whether it belongs to the deep learning related features at present according to the preset threshold. The cross-entropy loss function is used to improve the classification ability of the model, and the output of the feature extraction network can be used to judge whether the features belong to the same class or not. The calculation formula of the deep learning loss function is as follows: the loss function

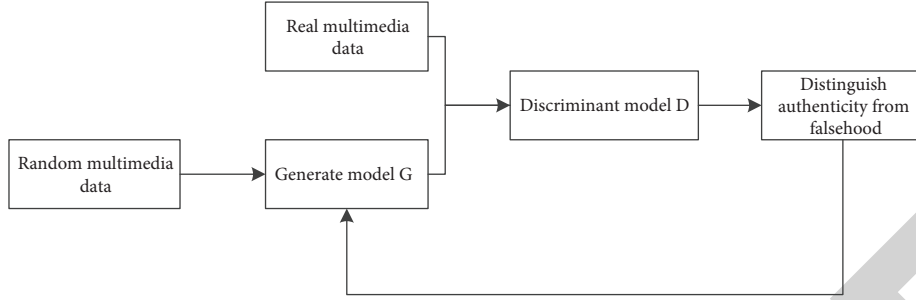


FIGURE 5: Deep learning GAN network model structure.

in deep learning can extract and classify multimedia pictures, videos, texts, and audio in the network through feature calculation. And the loss function has the ability to distinguish the structural model of neural network, and at the same time, it can strengthen the feature expression ability of the model. In the calculation of loss function, the relative loss of the deep learning neural network model in data processing can be obtained. Through the calculation of loss function, the application of the deep learning neural network model can be enhanced, and the loss degree in model application can be reduced, and the calculation accuracy can be increased.

$$P_{i,j} = \frac{e^{l_{i,j}}}{\sum_{K=1}^C e^{l_{i,k}}}, \quad (9)$$

$$\zeta_{ce} = -\frac{1}{N} \sum_{i=1}^N y_{i,j} (p_{i,j}).$$

In deep learning, because the recognition category is determined by the image ID, the loss function in this paper is also called ID loss function. Deep learning recognition can be regarded as a one-time learning task because the human eye ID in the test set has not appeared in the training set. In order to prevent the recognition model from overfitting and low generalization performance, label smoothing is used to smooth the cross-entropy loss function as follows:

$$L(\text{ID}) = -\frac{1}{N} \sum_{i=1}^N q_i \log(p_i) \begin{cases} q_i = 1 \\ q_i = 0 \end{cases}, \quad (10)$$

$$q_i = \begin{cases} \frac{\varepsilon}{N}, \\ 1 - \frac{N}{N-1} \varepsilon. \end{cases}$$

Due to the particularity of deep learning function recognition, in order to train the model to obtain a more differentiated feature representation, the expression formula of data-related feature function is as follows:

$$\xi_{\text{tripiet}} = \max(d_{(a,p)} - d_{(a,n)} + \text{margin}, 0). \quad (11)$$

There are three kinds of activation functions commonly used in deep learning neural networks, and the activation

functions mainly act on the output of neurons in the hidden layer, and the value change near 0 is obvious, and the value range is infinitely close to the coordinate axis, where the activation function calculation formula is as follows:

$$\text{Sigmoid}(x) = \frac{1}{1 + e^{-x}}. \quad (12)$$

The deep learning activation function can well deal with the phenomenon that the output of the function presents nonzero centralization, which will lead to the offset of the input data of the backward layer, thus reducing the speed of gradient descent. The specific calculation formula is as follows:

$$\text{Tan } h(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \quad (13)$$

The deep learning activation function refers to the linear flow whole function, which is essentially a ramp function. The whole function is divided into two parts. In the part where the input is less than 0 and the output of the function is 0. In the part where the input is greater than 0, the output value is equal to the input value. Compared with the first two activation functions, its convergence speed is faster and there is no saturation region, which can not only effectively solve the gradient change problem in the function but also be widely used in various depth networks. The specific calculation formula is as follows:

$$\text{ReLU}(x) = \begin{cases} x, & x \geq 0, \\ 0, & x \leq 0, \end{cases} \quad (14)$$

$$= \max(0, x).$$

From the three activation functions of deep learning, it can be concluded that the activation function is a linear flow whole function and has an algebraic formula. First, the activation function is mainly responsible for the output and input of neurons in the hidden layer of neural network. And the activation functions are divided into three categories; so, the output and input of neurons in different hidden layers of neural networks are responsible for different activation functions. According to the image of deep learning activation function, activation function is mainly a nonlinear function with a fixed operation value, which can operate all values because its function field is infinite.

The function images of the three activation functions of deep learning are shown in Figure 6.

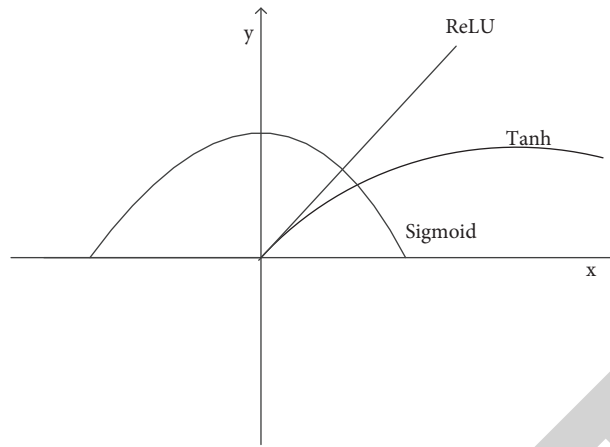


FIGURE 6: Image comparison of the deep learning activation function.

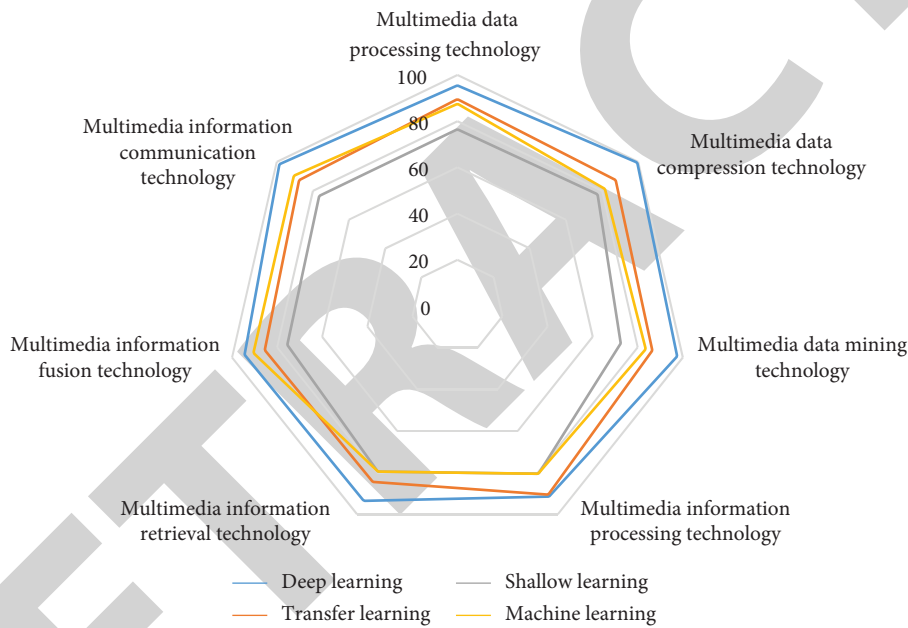


FIGURE 7: Performance comparison of deep learning in multimedia data processing.

4. Application Analysis and Research of Multimedia Data Processing Technology in Deep Learning

4.1. Performance Comparison of Deep Learning in Multimedia Data Processing. Many media file data are useful information; so, it is necessary to analyze this information for future research and utilization, which is a problem that needs in-depth study. However, with the in-depth application of learning methods, the traditional methods of obtaining effective information of multimedia data cannot be obtained accurately and quickly, which brings difficulties to data acquisition. Especially, multimedia data have a large amount of information, and there are much important information hidden in conflict information, which makes it difficult to mine. Under this background, how to realize the fast processing of massive multimedia data is an important

subject. The development of deep learning technology has led to the rapid growth of a large amount of information, including a large amount of multimedia information. How to quickly find the required information from the massive information has become an important issue. Therefore, this paper proposes a multimedia data processing method based on deep learning, which improves the accuracy of information retrieval by improving the feature extraction of multimedia data. However, with the development of the times, many media files have a large amount of data, but the requirements for data storage are quite different, and the real-time requirements are high, and the database structure changes greatly. In addition to processing and managing data, there are still a lot of problems. Therefore, multimedia information technology should have the following characteristics: the expression and processing of informal information such as video, audio, graphics, and multimedia; it

can control and reflect the space-time value and diversity of multimedia information, content-based query; ability to control versions and perform extended tasks; have network function; and differences of different multimedia information processing methods. Therefore, we analyze and compare the performance of deep learning in seven aspects of multimedia data processing as shown in Figure 7. From the radar chart in Figure 7, it can be seen that the overall research advantage line of deep learning is the outermost circle, which shows that the research advantage of deep learning is higher than that of transfer learning, and the research advantage of transfer learning is better than shallow learning and machine learning, while machine learning has the worst research advantage for multimedia data processing technology. Moreover, deep learning is beneficial to the processing research of multimedia data, and the research on multimedia information is also very prominent because deep learning has function-related algorithms and its neural network structure has great research advantages compared with multimedia data.

Deep learning is a part of machine learning, and it also realizes further learning in the field of machine learning. The deep learning method can automatically learn and obtain features according to input data, without time-consuming and labor-intensive manual selection, which greatly accelerates the completion of tasks. The depth of this method represents a series of continuous layers, while learning emphasizes learning from continuous layers, saving the data after each layer completes the operation in the weight of this layer, and realizing the transformation of each layer by parameter weight. The learning process of the model is to find a set of weight values on all layers of the model so that the output data corresponds to the target value. However, due to the large number of parameters, the one-to-one correspondence between the output value and the target value is not realistic; so, the evaluation index to measure the model error is needed. However, the research of deep learning needs modeling first. The main technical contents of deep learning are analyzed and explained as shown in Table 1. According to the content analysis in Table 1, deep learning mainly includes the random forest model, support vector machine, multi perceptron, convolution neural network, activation function, loss function, model optimization, and other technologies. Among them, stochastic forest model technology mainly solves the related data that need to be calculated by decision tree; support vector machine mainly deals with complex and regression data; and multi perceptron is a neural network model of perceptual data. The activation function and the loss function are used to analyze and calculate nonlinear mechanism and fitting data. Model optimization technology is used to optimize and train the parameters and functions of deep learning.

Deep reading learning has multiple application stacking layers, that is, the input of the next layer is equivalent to the data output description of the upper layer. This network has the same characteristics as the traditional neural network, that is, the hierarchical structure of the two networks, that is, the structure of the input-output network; there is no connection between adjacent nodes and there is no

relationship between the structures. Traditional neural networks map to selection function values. The training speed is slow and it is easy to adapt to inappropriate parameter changes. The selection network is deeply studied, and the hierarchical neural network is used to process the data complex. As a major information analysis tool, deep search has become an important practical field of digital research such as visual recognition and natural language processing. Through the research on the complex information value of deep learning multimedia data, the research on multimedia information service plays a very important role in the future development. The data processing parameters of deep learning are shown in Table 2.

4.2. Research on Multimedia Data in Deep Learning.

With the rapid development of deep learning technology, its application is becoming more and more extensive. Due to the large amount of multimedia data, the memory capacity and bandwidth of network communication are strictly required. Therefore, multimedia computing technology has become the key technology in multimedia, network communication, computer, and other application fields. This paper analyzes the application of some multimedia information processing algorithms in network maintenance and multimedia transmission. However, multimedia applications involve a variety of multimedia information, such as image, video, and audio information, which often require a variety of devices and flexible processing. In the image, we need to make some geometric adjustments, such as free scaling, translation, rotation and deformation, or smooth filtering and edge extraction. For video data, video can be superimposed or edited on the blue screen. In the case of audio data, filtering, reverberation, and noise control are also necessary. In addition, we also hope to query the data in the multimedia database. Multimedia information compression format further increases the difficulty of multimedia data processing. The background of this paper is the direct processing technology of deep multimedia information, which has obvious significance. Because multimedia technology greatly reduces the capacity of multimedia files, the processing efficiency should be greatly improved. Multimedia information processing technology has become a hot research topic in the multimedia field in the recent years. As an important research field of image and video processing, this technology has been widely recognized. The composition of the case base of multimedia data processing technology is shown in Table 3.

Multimedia refers to the expression forms of information, news, words, pictures, sounds, and videos. Especially, multimedia data are composed of various media, usually various media, such as pictures, pictures, news, audio, and video. The information of a certain media is called single media, and single media can also be considered as a special type of multimedia. Therefore, multimedia usually includes a single media (pure media), as well as multiple but combined media. There are many kinds of multimedia information, and different classification principles classify multimedia information differently. This

TABLE 1: Analysis and explanation of deep learning-related technologies.

Main techniques of deep learning	Analysis and explanation of related contents
Stochastic forest model	The main branch of deep learning is also the most representative decision tree model
Support vector machine	A learning algorithm for complex data processing and regression
Multi perceptron	A typical feed forward artificial neural network model
Convolution neural network	A deep neural network framework with convolution structure
Activation function	A core hyper of nonlinear fitting can directly affect the data processing ability of the deep learning model for nonlinear mechanism
Loss function	A deep learning model fits the data to be the state representation of the fitted data
Model optimization	Training optimization of complex set of parameters and functions of the deep learning model

TABLE 2: Related parameters of the deep learning network layer in multimedia data processing.

Deep learning network layers	Operation type	Filter (step size)	Output size
In-4a	Maximum pooling of data	3 * 3/2	5 * 5 * 515
In-4e	Maximum pooling of data	3 * 4/3	6 * 6 * 512
In-SB	Maximum pooling of data	3 * 3/2	7 * 7 * 512
Ip-4a	Data convolution	3 * 3/1	6 * 6 * 1024
Ip-4e	Data convolution	3 * 3/1	4 * 4 * 256
Ip-5e	Data convolution	3 * 2/3	5 * 5 * 25
Merging layer	Data tandem	4 * 4/1	4 * 4 * 768
Convolution layer	Data convolution	1 * 1 * 2	6 * 6 * 128
Pool layer	Data summation pooling	1 * 1/5	1 * 1 * 128
Depth characteristic layer	Data characterization	150	1 * 1 * 150

TABLE 3: Related components of multimedia data processing technology.

Case base names	Typical case	Emphasis
Multimedia data representation	Face image	Data acquisition display
Digital image processing	Face detection	Human-computer interaction data
Speech signal processing	Speech recognition in human-computer interaction	Education and training
Multimedia data compression	Compressed sensing technology and its application	Data storage
Multimedia data retrieval	Feature extraction of multimedia data	Multimedia data transmission retrieval
Multimedia information retrieval	Multimedia data matching	Multimedia image retrieval
Multimedia fusion data processing	Multi human-computer interaction multimedia data	Family entertainment

paper focuses on the proportion of multimedia data types. Through the experimental analysis of deep learning, it is concluded that the proportion of multimedia data types is as shown in Figure 8.

Compared with traditional characters, multimedia data have its own characteristics in data type. Large amount of data is an important feature of multimedia data, and it also makes multimedia data processing technology quite difficult. Multimedia data are usually composed of many different types of unstructured single media data. Compared with their unique single media, multimedia has the characteristics of combining sound, text, and image, and the data amount of multimedia is several times that of single media. Interface is an important feature of multimedia technology which means that people can process and generate multimedia information. Video and audio are based on time information and have time characteristics. From the characteristics of multimedia, we can see that multimedia data have many characteristics, and the characteristics of its data have relevant research and analysis significance for deep learning. A brief description of the relevant characteristics of

multimedia data is shown in Table 4. The characteristics of multimedia data are mainly quantitative (the amount of data is very large), integration (there are various types of media data that are combined), interactivity (multimedia data can be processed and created), real-time (multimedia data have time characteristics), nonstructural (data have different coding methods and different processing methods, etc.), dynamic (with various characteristics of modifying data), nonlinear (with very flexible change methods), and controllable (multimedia data have controllability and can be expressed in the way of demand).

Most of the multimedia data are stored in binary form, and multimedia data have different structure and coding methods. Therefore, with the deep study of the characteristics of Indian multimedia data types, the characteristics of multimedia data are very prominent, which are more conducive to the development of multimedia data processing technology. We set the overall effect of studying multimedia data features to 100 points, and 90–100 points are good for studying features; 80–90 is the best research feature; 70–80 is the average research effect; 60–70 is a poor research feature;

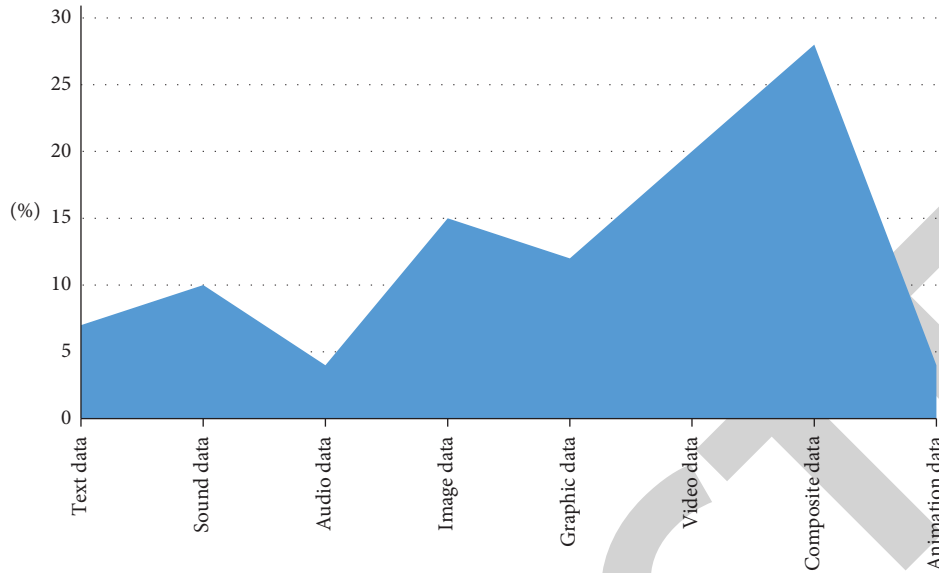


FIGURE 8: Relevant proportion of multimedia data types.

TABLE 4: Correlation analysis and explanation of multimedia data characteristics.

Multimedia data characteristics	Analysis and interpretation of multimedia related features
Quantitative	The amount of multimedia data are very large
Integration	Multimedia data are mainly composed of a variety of different types of media
Interactivity	Multimedia data can be processed and created
Real-time	Multimedia is a kind of time-based media with time characteristics
Nonstructural	Multimedia data have different coding methods and their respective processing methods and structures
Dynamics	Multimedia data have the characteristics of random combination/addition/deletion
Nonlinear	Multimedia data are presented in a flexible and varied manner
Control	Multimedia data can be controlled and expressed in the form of demand

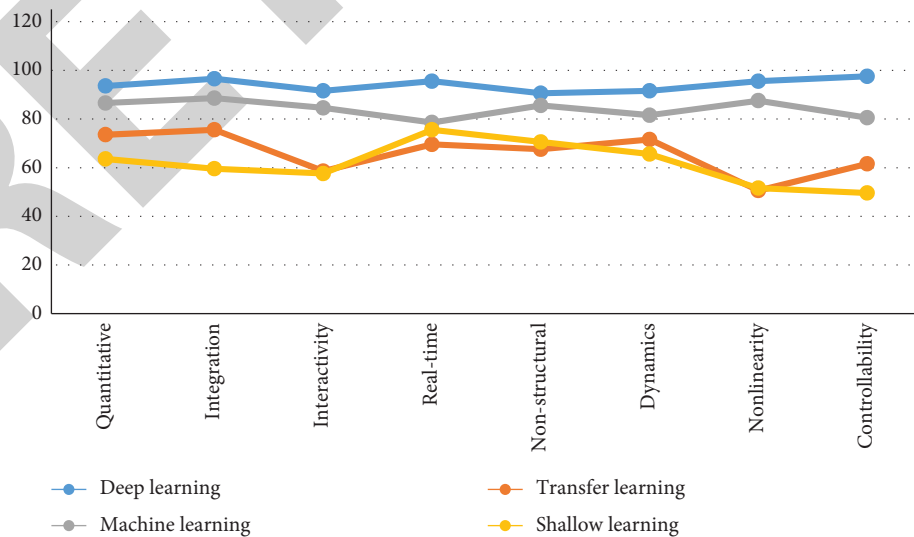


FIGURE 9: Correlation analysis and comparison of multimedia data features by deep learning.

and 60 for poor research results. The data features of multimedia are easier to show under the analysis and comparison of deep learning, compared with the depth of transfer learning, machine learning, shallow learning, and so on. The relevant comparison of deep learning for multimedia feature research is shown in Figure 9. From Figure 9, it can be analyzed that there are various features of multimedia data, but the features of multimedia data analysis in deep learning are higher than those in transfer learning, machine learning, and shallow learning. It shows that deep learning is very suitable for feature analysis of multimedia data. In the analysis of the characteristics of multimedia data, such as quantification, integration, interactivity, real-time, unstructured, dynamic, nonlinear, and controllable, the research values of deep learning are all between 80 and 100, and the research effect is obviously higher than other research methods.

4.3. Application Correlation Analysis of Multimedia Data Processing Technology. With the continuous development of multimedia data processing technology, multimedia data processing technology has also been the corresponding progress and promotion. In people's current work and life almost everywhere, we can see the application of multimedia data processing technology. The efficiency of people's work and study has been greatly improved because of the emergence and application of multimedia data processing technology. Among them, the application of multimedia data processing technology mainly refers to the integration of text images, audio, and video elements so that relevant information can be better displayed in front of people. The significance of using multimedia technology lies in expanding the application range of computer technology accurately. Multimedia technology also makes the data interface more friendly, and professionals can control the use of multimedia devices in a short time. With the emergence of multimedia information technology, audio and video technology and communication technology are closely linked in three different fields, which lay a good foundation for the rapid development of information technology. The application comparison of multimedia data processing technology is shown in Figure 10. From the analysis in Figure 10, it can be concluded that the related applications of multimedia data processing technology in image application field, text application field, audio application field, video application field, and composite application field are slightly higher than other multimedia processing technologies. Therefore, multimedia data processing technology is the core technology in the multimedia technology center, in which multimedia protective gear compression technology and multimedia information related technology are auxiliary technologies. Generally speaking, more than these technologies, multimedia data processing technology has stronger ability to analyze and process data.

Multimedia information processing technology is divided into video processing, image processing, audio data processing, audio data processing, network data processing, data processing, and composite data processing. In addition, multimedia information technology is also applied in several important fields. In concrete application, multimedia technology can be divided into communication system realization, editing system realization, industrial application, medical imaging system, and education. To a certain extent, multimedia information technology provides convenience for modern life and improves the ability of information transmission. In the industry, multimedia mainly develops the market and trains talents through multimedia teaching. This form not only greatly reduces the production cost but also promotes the application of surgical medical imaging system in the medical market and becomes an important symbol of the development of medical industry. Using multimedia information processing technology can improve the effect of medical image analysis and processing, thus achieving remarkable results. The application of multimedia courseware in language and writing makes teachers and students deepen their understanding of knowledge through film and television, which has achieved good results and improved students' learning efficiency and level. Among them, the application of multimedia data processing technology in various fields is shown in Figure 11.

Multimedia contains a variety of media data, which mainly includes multimedia format data, multimedia spatial data, multimedia content data, multimedia time data, multimedia version data, multimedia network data, multimedia open data, and multimedia thing data. In order to study the performance of multimedia data processing technology for related multimedia data processing, we compare multimedia data mining technology with multimedia data compression technology for experimental analysis. If the multimedia data processing is completed by multimedia data mining technology, multimedia data compression technology, and multimedia data processing technology, among them, through deep learning experiment and data analysis, it is concluded that multimedia data processing technology is superior to the other two core technologies in terms of multimedia related data processing ability. The performance comparison of multimedia data processing technology is shown in Figure 12. By analyzing Figure 12, we can find that there are various types of multimedia data, including data, spatial data, content data, time data, version data, network data, open data, and transaction data. Compared with these multimedia data, multimedia data processing technology, multimedia data mining technology, and multimedia data compression technology all have the ability to process these data. But by analyzing the data in the graph, it can be concluded that the multimedia data processing technology is much better for these data processing effects.

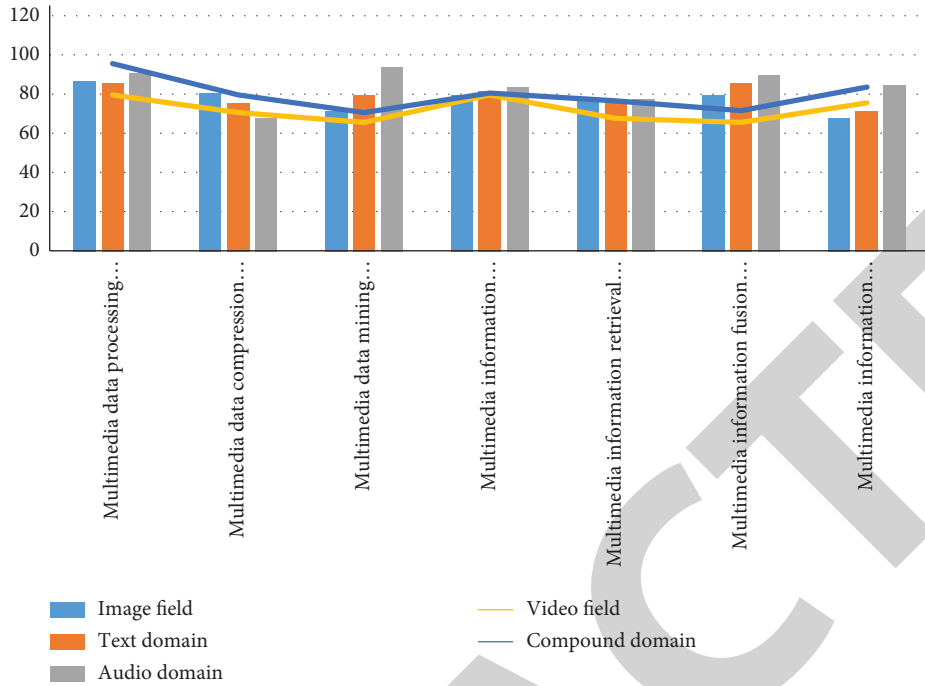


FIGURE 10: Comparison of application fields of multimedia data processing technology and related technologies.

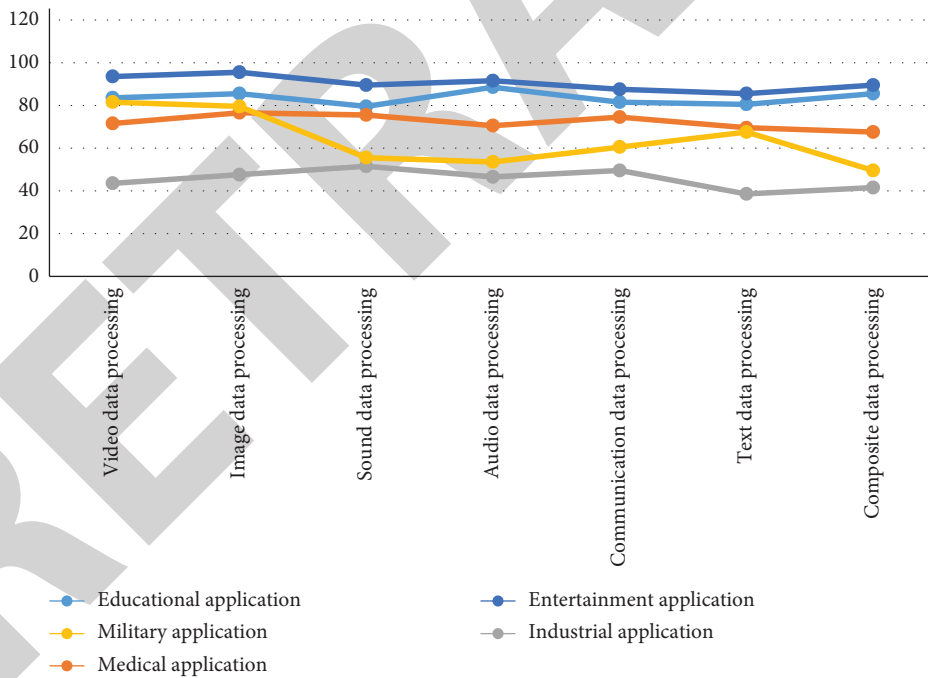


FIGURE 11: Application proportion of core technologies of multimedia data processing.

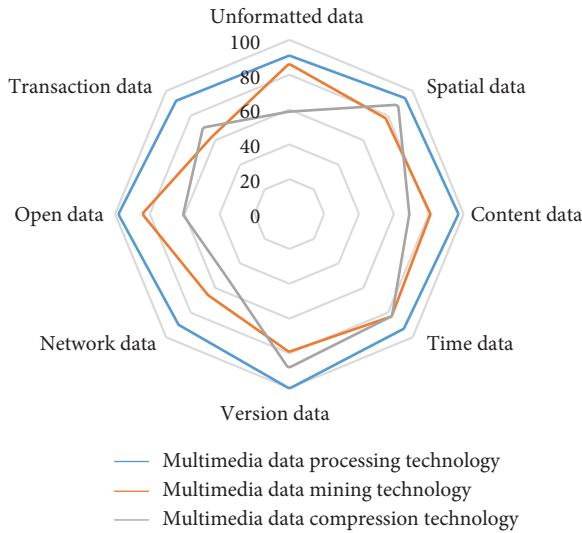


FIGURE 12: Performance comparison of multimedia data processing technology.

5. Concluding Remarks

With the popularity of the Internet and the application of multimedia in modern life and the development of database, the amount of multimedia data has gradually increased. Based on deep learning, this paper makes an in-depth analysis and research on the application of multimedia data processing technology. Data analysis shows that deep learning has comparative advantages in all major research fields, especially in multimedia research fields. Therefore, under the research of deep learning, we conclude that multimedia data are unstable and have many characteristics, and different multimedia data have different characteristics, mainly manifested as integration and nonstructural characteristics. Moreover, the composition of multimedia data processing technology is complex and its data processing methods are also various. Multimedia data processing technology can classify data according to different characteristics of data, which is beneficial to the expression of multimedia to be more accurate and comprehensive. The main manifestation of multimedia is to display through image and audio, so multimedia data processing technology is also very important for image and audio data processing. However, the performance of multimedia data processing technology is expensive, so we need to improve the processing efficiency of multimedia data processing technology. And an in-depth study of multimedia data processing technology to improve its shortcomings is needed for the powerful multimedia to provide better data processing technology.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the Foundation of Excellent Young-Backbone Teacher of Colleges and Universities in Henan Province under Grant 2019GGJS182 and in part by the Key Scientific Research Project of Henan Colleges and Universities under Grant 21B120001.

References

- [1] H. Wang, N. Wang, and D. Y. Yeung, "Collaborative deep learning for recommender systems," in *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, vol. 26, no. 4, pp. 1–11, Singapore, August 2019.
- [2] Z. J. Sun, L. Xue, and X. U. Yang-Ming, "Overview of deep learning," *Application Research of Computers*, vol. 27, no. 1, pp. 15–22, 2021.
- [3] R. Collobert, "Deep learning for efficient discriminative parsing," in *Proceedings of the 14th international conference on artificial intelligence and statistics*, vol. 49, no. 2, pp. 764–626, Lauderdale, FL, USA, April 2020.
- [4] Y. Bengio, G. Guyon, and V. Dror, "Deep learning of representations for unsupervised and transfer learning," *Workshop on Unsupervised and Transfer Learning*, vol. 21, no. 4, pp. 154–184, 2021.
- [5] V. Dumoulin and F. Visin, "A guide to convolution arithmetic for deep learning," no. 2, pp. 271–624, 2020.
- [6] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Transactions on Cognitive Communications and Networking*, vol. 11, no. 4, pp. 2715–2614, 2021.
- [7] J. Glassy and D. Cubanski, "Geodigital multimedia data processing system and method," no. 1, Google Patents, US Patent App. 10/980,577, pp. 219–272, 2020.
- [8] J. Molina, J. M. Mora-Merchan, and J. Barbancho, "Multimedia data processing and delivery in wireless sensor networks," *Intech*, vol. 50, no. 1, pp. 14–27, 2020.
- [9] S. Bhattacharyya and U. Maulik, *Soft Computing for Image and Multimedia Data Processing*, Springer, Berlin Heidelberg, 2019.
- [10] H. Frank, "Multimedia data maintaining and processing device," *Alexandria Engineering Journal*, vol. 21, no. 2, pp. 1761–1524, 2021.
- [11] Y. S. Jeong, M. L. Shyu, G. Xu, and R. R. Wagner, "Guest editorial: advanced technologies and services for multimedia big data processing," *Multimedia Tools and Applications*, vol. 74, no. 10, pp. 3413–3418, 2015.
- [12] T. Kartbayev, R. Uskenbayeva, and M. Mussaif, "Effective algorithms of multimedia data processing," *dspace*, vol. 9, no. 4, pp. 243–447, 2020.
- [13] X. Zhao, "A study of the 0/1 knapsack problem in the application of the dynamic self-adapting multimedia data processing method," *Computer Engineering and Science*, vol. 33, no. 8, pp. 154–157, 2021.
- [14] M. Usman, M. A. Jan, X. He, and J. Chen, "A survey on big multimedia data processing and management in smart cities," *ACM Computing Surveys*, vol. 52, no. 3, pp. 1–29, 2019.
- [15] X. Hui and J. Shi, "The study of compressed domain multimedia data processing," *Journal of Image and Graphics*, vol. 21, no. 2, pp. 1463–1721, 2021.