

## Research Article

# Classification and Retrieval of Multimedia Book Cloud Resources Based upon Hash Algorithm

Li Shen 

Shandong Women's University, Jinan, Shandong 250300, China

Correspondence should be addressed to Li Shen; 22011@sdwu.edu.cn

Received 28 January 2022; Accepted 21 February 2022; Published 17 March 2022

Academic Editor: Hye-jin Kim

Copyright © 2022 Li Shen. 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.

Classified retrieval is one of the important factors of library management. It can help readers and administrators quickly find the required materials and better enjoy the convenience of modern network. Through data mapping, hash algorithm can directly convert high-dimensional data into low-dimensional data and store it in Hamming space, which greatly reduces storage consumption and increases computational convenience. For the multimedia retrieval and classification in the era of big data, hash algorithm is right to the point. This paper is aimed at studying the classification and retrieval method of multimedia book cloud resources, extracting different classification categories, determining classification items; cloud resources contain a large amount of image text information, and according to the characteristics of the image text, the features can be well extracted for classification and retrieval, find the optimal book retrieval and classification method, and realize the classification and retrieval of multimedia book cloud resources with the help of hash algorithm. The experimental results show that the hash algorithm can help the classification and retrieval of multimedia books, optimize the classification of multimedia book cloud resources, and improve the efficiency by 20% and user satisfaction by 30%.

## 1. Introduction

In the wave of big data, cloud computing-related industries have developed rapidly and have become an indispensable part of people's life. Cloud computing-related industries have attracted the attention of more experts and scholars, and relevant research has been carried out in turn. Relying on big data, multimedia cloud resources have a great impact on the development of film and television, short video, and other multimedia industries. However, resource supply and task scheduling are also facing new challenges because of the massive data, high competitiveness, and immediacy of multimedia cloud. Effective classified retrieval of cloud resources can improve efficiency and resource utilization on the one hand and is of great significance to promote the stable development of multimedia cloud and ensure user experience on the other hand. Traditional keyword research methods cannot meet the information needs of users. In order to make up for the deficiency of keyword research methods, this paper proposes a content-based video research method. This search method can effectively improve the

search accuracy. Therefore, this paper studies the search of classified multimedia book cloud resources and constructs a book search framework based on the content of multimedia book cloud resources.

Combined with the relevant theories of hash algorithm, this paper puts forward a book search method based on the integration of multimedia book cloud resources, hoping to provide a valuable reference for the research of digital library. For optimizing book retrieval methods, experts and scholars at home and abroad have carried out in-depth research and made some achievements. Gardezi et al. proposed a new method of classification using time series analysis. In particular, they use dynamic time warping as a similarity measure to classify the region of interest in the picture as normal or abnormal. This method is particularly attractive for image analysis. The current research concludes that changing the size of the classified image of the region of interest and the restrictions on the distorted path search criteria will not affect the performance, because this method produces good classification results and reduces the computational complexity [1]. Zhang and Jiang proposed a new

tensor-based logistic regression algorithm to complete multimedia classification through Tucker decomposition. In order to strengthen the classification process, the regularization term adopts  $F$ -norm and establishes logistic Tucker regression model to effectively extract principal components from tensors, so as to reduce the input dimension and improve the efficiency of multimedia classification [2]. Bissotti et al. make a comparative study on six methods of texture 3D model retrieval and classification and studies the analysis methods of specific categories of geometric and texture deformation. By constructing a set of 572 synthetic texture mesh models, each class includes multiple texture and geometric modifications of a small group of empty models. The results show a challenging but vivid scene and reveal interesting insights on how to process texture information according to different methods [3]. Narducci et al. study the problem of using appropriate algorithms for TV program classification and retrieval in the context of establishing personal channels. They propose a new feature generation technology, which uses additional features extracted from Wikipedia to enrich the text program description. This paper introduces the concept of personal channel of EPG, finds and classifies them by using program types and short text descriptions of programs, and provides users with potentially interesting programs and videos [4]. Adhikary et al. have developed a quality of service- (QoS-) aware cloud resource management system to reduce energy consumption and improve resource utilization through various multimedia social applications. In order to minimize VM creation time, it allows the recycling of virtual machine (VM) resources for user requests with similar resource requirements, so as to minimize underutilization or overutilization of resources and improve user satisfaction. Recent research shows that gratifying progress has been made in solving classification problems [5]. Liu et al. proposed a new loss function to take advantage of the relationship between categories and classifiers. In addition, bipolar relationship (BR) diagrams are used to develop a common form for various relationships. The bipolar graph is automatically learned to reexperience constraints that may occur during cost minimization. A large number of experiments on three benchmarks with various assumptions and graphs show that this method can provide significant performance improvement by learning together from BR graphs and assumptions, especially in small training data set scenes suffering from serious overfitting problems [6]. Singaravelan et al. extract information from multiple texts on the same topic through multiple document summaries. The resulting summary report enables individual users (such as professional information consumers) to quickly become familiar with the information contained in a large number of documents. Their clustering-based importance ranking method (CBRS) summarizes multidocuments with semantic significance. In this way, it will produce good results when clustering and sorting documents. It can be used as the clustering result to improve or refine the sentence ranking result. The effectiveness of this method is proved by clustering quality analysis and summary evaluation of simulated data sets [7]. Al-Akashi and Inkpen design a real-time network search

method. Aggregate several web search algorithms during query to adjust the relevance of search results. They learned a context-aware delegation algorithm that allows the selection of the best real-time algorithm for each query request. The evaluation shows that the proposed method is superior to the traditional model. It is highly correlated with the recently searched query, has the same performance, and can adapt to the shortcomings faced by other algorithms [8].

By consulting relevant literature, this paper understands the shortcomings and advantages of traditional data cloud database classification and retrieval. This paper introduces the hash algorithm and constructs the classified retrieval model of multimedia book cloud resources based on the hash algorithm, so as to better realize the purpose of multimedia book cloud resource retrieval. Its advantages are as follows: (1) a detailed and specific introduction to the features of images and texts and how to extract features, (2) the Lagrangian algorithm is introduced to better optimize the image text retrieval method, and (3) introduced multiple evaluation indicators such as accurate search and complete search and retrieval time, and the structure is more accurate.

## 2. Multimedia Book Retrieval Method and Hash Algorithm

### 2.1. Hash Algorithm

*2.1.1. Approximate Nearest Neighbor Retrieval Based on Hash Method.* In the early stage of the development of retrieval task, a typical and effective method is nearest neighbor retrieval, that is, by calculating the distance between multimedia data and searching from the data set to obtain the data closest to the query data. This method is simple in principle and easy to implement. However, with the massive growth of multimedia data, if the nearest neighbor retrieval method is applied to massive data sets, the computational complexity will become great and the retrieval efficiency will be affected. Therefore, large-scale multimedia retrieval tasks have attracted more and more attention. Researchers have proposed a more effective alternative method, namely, approximate nearest neighbor (ANN) search. The most typical one is the hash method.

The retrieval process of the hash method can be summarized as follows: Firstly, the hash function and the hash code of the retrieval database are obtained through the designed algorithm model and training data (generally composed of "0" and "1" strings). Then, the hash function is applied to the data to be queried to generate the hash code corresponding to the data to be queried. In the process of retrieval, the similarity between the data to be queried and the corresponding hash code of the database is calculated through XOR operation. That is, the smaller the Hamming distance, the greater the degree of similarity between the two samples. On the contrary, the smaller the similarity between the two samples, the higher the similarity will be returned to the user.

*2.1.2. Image Retrieval Based on Graph Regular Hash.* Graph-based hashing technology is an important method in large-

scale image retrieval hashing technology. This kind of method can make use of the flow structure in the data set and maintain the nearest neighbor relationship between data through Laplace regularization [1], for example, spectral hash (SH), anchor graph hash (AGH), scalable graph hash (SGH), large-scale spectral rotation hash (lghsr), and discrete graph hash (DGH).

SH is improved on the basis of the semantic hash method, which considers that the problem of generating optimal hash code using a data set is similar to the graph segmentation problem (NP hard problem). In order to solve this problem, SH first performs spectral operation on the original high-dimensional space samples, then calculates the eigenvector of Laplacian graph through relaxation constraints, selects the eigenvector quantum set corresponding to the eigenvalue greater than a certain threshold, and binarizes it to obtain hash code. The process is as follows:

Suppose that the data set  $M = \{m_1, m_2, \dots, m_n\} \in Z^{k \times b}$  contains  $k$  samples, and the characteristic dimension of each sample is  $b$ . The purpose of graph hash is to map each sample into a  $D$ -bit hash code  $h_x \in \{-1, 1\}^{1 \times d}$ , where  $d < b$ . SH expects to represent the hash code of similar data by as few bits as possible, and the probability of 0 or 1 of each hash code bit is the same. And the bits are independent of each other, so the target formula takes into account both balance and independence. SH formula model is calculated as follows:

$$\min_F \sum_{xy} Q_{xy} \|h_x - h_y\|^2 \quad (1)$$

$$\text{s.t. } h_x \in \{-1, 1\}^l$$

$$\sum_x h_x = 0 \quad (2)$$

$$\frac{1}{k} \sum_x h_x h_x^T = A,$$

where  $Q_{xy}$  represents the similarity matrix between samples,  $\sum_x h_x = 0$  represents the balance of hash codes, and  $1/k \sum_x h_x h_x^T = A$  represents the independence between bits.

Since the solution is NP hard, SH relaxes the constraints and uses spectral analysis technology. The objective function after relaxation is as follows:

$$\min_F Td(H^T L H) \quad (3)$$

$$\text{s.t. } H^T \mathbf{1} = 0$$

$$H^T H = kA_d \quad (4)$$

$$H \in \{-1, 1\}^{k \times d},$$

$$L = C - Q, \quad (5)$$

$$C_{xx} = \sum_y Q_{xy}, \quad (6)$$

where  $h$  represents the graph embedding matrix of  $K$  samples,  $l$  represents the Laplace matrix, and  $C$  represents

the diagonal matrix. It can be seen that the spectral hash method introduces the Laplace matrix  $L$  and eliminates the discrete constraints. The flow pattern learning method can be used to obtain the eigenvector of the Laplace matrix. Finally, the hash code is obtained by thresholding the eigenvector by the symbolic function.

Although SH can achieve good performance, when the number of data set samples is large, the calculation cost of calculating  $l$  will increase greatly. The AGH method proposes a new optimization scheme to solve this problem; that is, the clustering center obtained by  $K$ -means is regarded as an anchor point. When the number of anchor points is large enough, AGH uses the anchor point data and sample data to construct the anchor point diagram and then constructs a nonnegative, sparse, and low-rank adjacency matrix to approximately replace the original adjacency matrix, so as to greatly reduce the computational overhead. The details are as follows:

Assuming that the data set  $M$  contains  $k$  sample data, first use the  $k$ -means method to obtain  $g$  ( $g \ll k$ ) cluster centers (anchor points)  $P = \{p_1, p_2, \dots, p_n\} \in Z^{k \times b}$  on the data set  $M$  and then calculate the similarity matrix  $X$  between the sample and the anchor points:

$$X = \begin{cases} \frac{K(x_a, p_b)}{\sum_{\tilde{b} \in \langle a \rangle} K(x_a, p_{\tilde{b}})}, & \forall b \in \langle a \rangle, \\ 0, & \text{other,} \end{cases} \quad (7)$$

$$K(x_a, p_b) = \exp\left(\frac{-\|x_a - p_b\|^2}{2\alpha^2}\right). \quad (8)$$

Based on the similarity matrix  $X$ , the normalized similarity matrix between samples can be obtained:

$$B_{a,b} = \sum_{d=1}^k W_{ad} W_{bd} = W_a W_b^T, \quad (9)$$

$$B = W \Lambda^{-1} W^T, \quad (10)$$

$$\Lambda = \text{diag}(W^T \mathbf{1}), \quad W \in G^{k \times b}, \quad (11)$$

$$L = A - B. \quad (12)$$

Because the dimension of the constructed anchor graph is  $k \times b$ , the dimension of the original adjacency matrix is  $k \times k$ , and  $b$  is far less than  $k$ ; this method can greatly reduce the time complexity.

The optimization problems of AGH and SH are the same, except that AGH further relaxes the discrete constraints; that is, remove the equilibrium constraint  $W^T \mathbf{1} = 0$ :

$$\min_F Td(H^T B H) \quad (13)$$

$$\text{s.t. } H^T \mathbf{1} = 0$$

$$H^T H = kA_d. \quad (14)$$

Although the relaxation solution of the formula can be constructed by calculating the eigenvectors corresponding to the first  $D$  eigenvalues of  $B$ , the direct eigenvalue decomposition of  $B$  will bring large computational overhead. Therefore, the construction matrix is calculated as follows:

$$V = \Lambda^{-1} W^T W \Lambda^{-1/2} \in U^{d \times d}. \quad (15)$$

The singular value decomposition of  $V$  is selected to obtain the following:

$$V = R \sum R^T. \quad (16)$$

The solution of hash  $H$  is calculated as follows:

$$H = \sqrt{k} W \Lambda^{-1/2} R \sum - \frac{1}{2}. \quad (17)$$

## 2.2. Multimedia Book Retrieval Method

**2.2.1. Multimedia Cloud.** Multimedia cloud is a cloud resource developed on the basis of cloud computing, which can provide multimedia services such as social networking, Internet, radio, and television. Nowadays, with the help of multimedia cloud, people can understand information consultation all over the country, which also provides opportunities for the development of multimedia industry. Multimedia cloud has become a research hotspot, and its application is affecting the production and life of millions of people. Figure 1 shows the function of multimedia cloud [9].

**2.2.2. Hierarchy of Multimedia Cloud System.** According to different service types, cloud computing services are generally divided into three levels: IAAs, PAAS, and SaaS. The cloud-based multimedia system can also be understood from the perspective of multilevel services, as shown in Figure 2. IaaS refers to infrastructure as a service, which provides external hardware services. PaaS refers to platform as a service, which can provide a platform service for development and design. SaaS refers to software as a service, which can provide software operation services.

IAAs layer uses virtualization technology to provide it, storage, network transmission, and other infrastructure services, while PAAS services can run in cloud infrastructure or ICT origin facilities, or a mixture of the two resources. Similarly, SaaS layer multimedia applications can use the media services provided by PAAS layer. Using media services through cloud API (application programming interface), mix the two services or run on the server architecture. The layered perspective provides a conceptual level for understanding cloud multimedia systems [10].

**2.2.3. Framework of Content-Based Image Retrieval System.** At present, content-based image retrieval technology has been relatively mature. Many image retrieval systems have been established in both academic and commercial fields, and satisfactory results have been achieved. By reading a large number of relevant literature, it can be seen that a complete content-based image retrieval system should have three

core modules: image feature extraction, retrieval, and user relevance feedback [11]. The system framework is shown in Figure 3.

The process of extracting image features is also called the process of representing image features. At this time, the system should read all the files in the image database, extract the visual features from the image through the algorithm, and store the features in the image database or index file. In the search stage, the same feature extraction algorithm is used to obtain the feature vector of the image to be queried, calculate the similarity between the vector and all vectors in the feature library, sort the images according to the similarity, and finally return a group of similar images to the user. Semantic gap means that people usually judge the similarity of images not based on the similarity of the low-level visual features of the image, but based on the semantic understanding of the objects or events described by the image, and there is a wide semantic gap in image retrieval. Then, users can provide relevant feedback according to the search results, further optimize the search results, and narrow the “semantic gap” [12].

### 2.2.4. Key Technologies of Image Retrieval

**(1) Color Features.** Color feature is the first feature applied to image search system. The extraction process is simple and clear, is not easy to be affected by the change of size and viewing angle, and has good stability [13]. Until now, color is still the most commonly used method to represent features in image search system. Before extracting color features, we need to select the appropriate color space model and feature extraction algorithm. Color space model is a mathematical model that summarizes color information. Common color models include RGB, HSV, CMY, and lab [14]. Among them, RGB is the most commonly used model in electronic display devices. It can represent any color by layering and mixing three colors in different proportions. Its disadvantage is that the model is incompatible with human visual senses and very abstract. In contrast, HSV color space is more intuitive for people. HSV model is widely used in CBIR system because it accurately reflects the understanding of color through three attributes: brightness, saturation, and chroma. CMY is the abbreviation of cyan, magenta (magenta), and yellow (yellow), plus black (black), that is, CMYK subtractive color mixing mode. The color produced by this method is called subtractive color because it reduces the reflected light required for the visual system to recognize the color. Lab color space (English: lab color space) is a color-opposite space, with dimension  $L$  for brightness and  $a$  and  $b$  for color opposing dimensions, based on nonlinearly compressed CIE XYZ color space coordinates. The main algorithm is shown in Figure 4.

**(2) Texture Feature.** Texture feature is the basic attribute of object surface, which is used to represent the transformation or repetition of image in gray or color space. It is not affected by color or brightness and can describe the physical characteristics of the target or person in the image and the relationship between the target and the surrounding space [15]. The

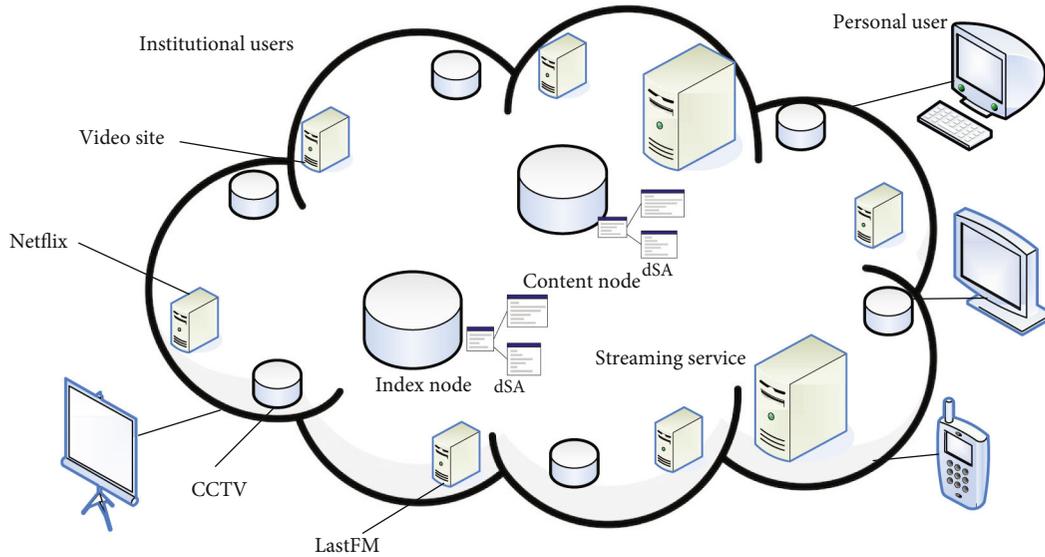


FIGURE 1: Multimedia cloud resource usage.

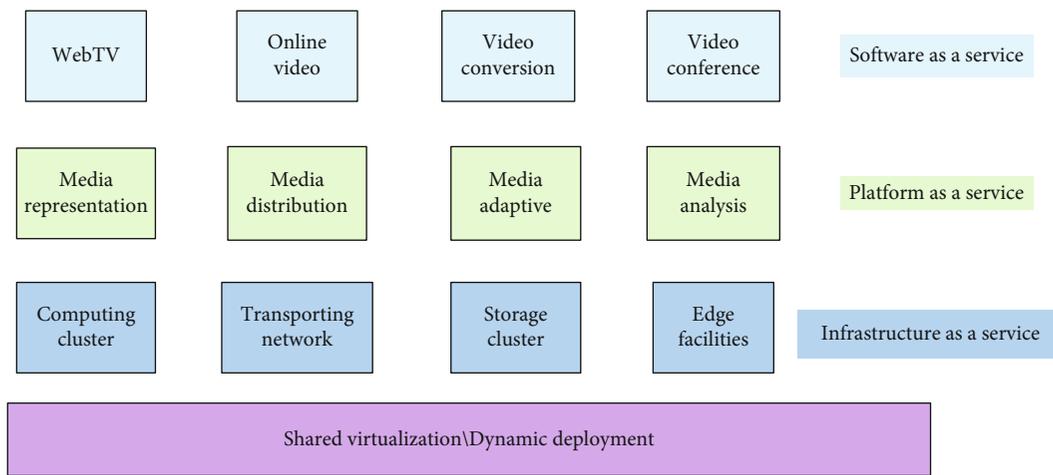


FIGURE 2: Hierarchical structure of multimedia cloud system.

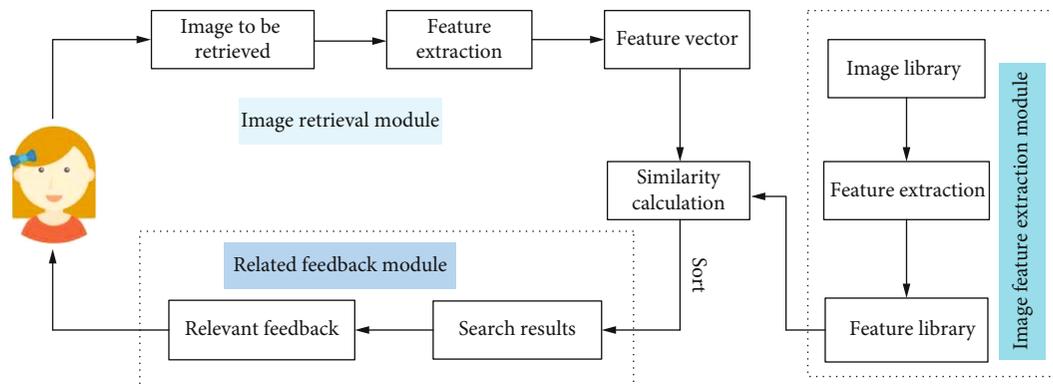


FIGURE 3: Content-based image retrieval system framework.

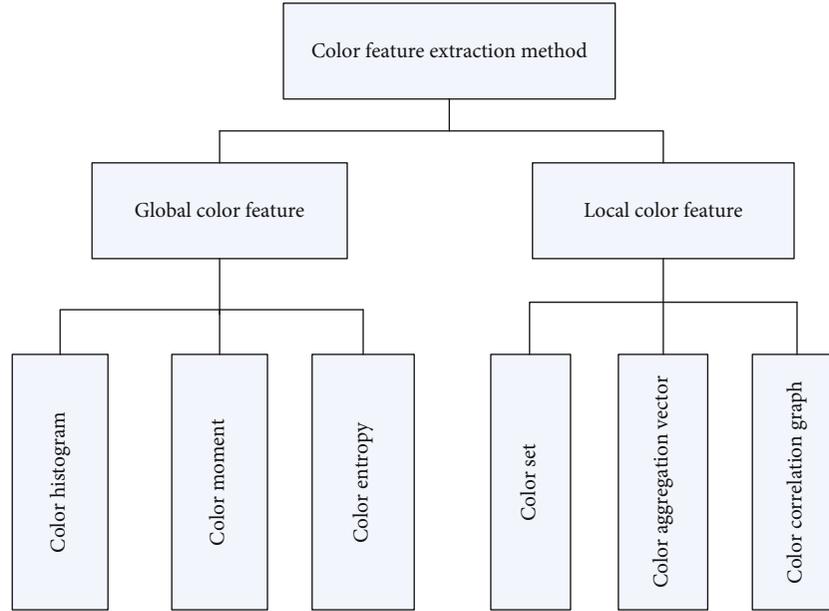


FIGURE 4: Color feature extraction method.

difference of texture can be used to recognize and classify images. Texture description algorithms can be divided into structural methods, model methods, statistical methods, signal processing, and other types. Each type contains a variety of texture description algorithms. These methods can be roughly divided into two categories: statistical analysis methods and structural analysis methods. Statistical texture analysis looks for digital features that characterize the texture, and uses these features or combined with other nontexture features to classify regions (rather than individual pixels) in the image. Structural texture analysis studies the primitives that make up the texture and their arrangement rules.

(3) *Spatial Features*. In the extraction direction, in addition to considering the shape, texture, and other features of the image, the position of the object in the image and the spatial relationship between objects are also very important. Therefore, the method of image extraction proposes spatial features. This method is based on the spatial position of multiple segmented objects in the image, such as the relationship between the distance and size or relative direction of the object, as well as the relative situation of up, down, left, and right. It effectively compensates for defects that cannot be described in the spatial information of the image for other features [16].

(4) *Semantic Features*. In addition to color, texture, and shape, images also have high-level semantic features. The understanding of images by human thinking comes not only from one aspect of low-level features. Therefore, when detecting images, we should consider the rich expression connotation of human language, so that the final result of search is close to human's actual intention. Due to the difference between visual level and semantic level, there is a semantic gap between low-level features and high-level fea-

tures. A method of reconstructing text image is proposed to improve the relevance of semantic search. By designing a two-level scoring system that automatically identifies queries and concepts, we can solve the problem of how to map text queries based on semantic analysis at a higher level, and significantly improve the semantic analysis of search results [17]. In addition, the database can be restored using low-level functions, and then, the high-level functions can be restored by comparing the same identifiers of different image layers, so as to reduce the semantic gap between them and the functions of high-level and low-level features.

2.3. *Technical Overview of Text Classification Methods*. The classification of Chinese texts mainly includes two stages [18]. The first step is the training stage, that is, the learning stage. The text classifier is generated by processing the known classified text set in this step. The second step is the classification step. Select the unclassified text samples and send them to the classifier after text preprocessing. The classifier automatically determines which category the text should belong to. The main steps of text classification process include the following: preprocessing, feature extraction, text formation, and obtaining classification results.

Text preprocessing is to convert unstructured data that cannot be processed directly by computer into structured data that can be processed directly by computer. Text preprocessing mainly includes extracting information, selecting text features, word segmentation, and deactivating words.

Information extraction refers to extracting the topic, title, and other information of the article we are interested in from the original HTML page crawled down by the web crawler. This information is meaningful for subsequent classification. Text feature selection means that we classify documents according to basic elements such as words, phrases, sentences, and even chapters. The more information contained in the element, the more difficult it is to process.

Therefore, according to the characteristics of Chinese, we choose words as text features. Chinese word segmentation is a key step in Chinese text processing, which will be described in detail later. Inactivated words refer to those words that appear more frequently in the text, but have no guiding significance for text classification, for example, modal particles such as “ah” and “Na,” quantifiers such as “once” and “Yi,” and pronouns such as “I” and “you.” These words should be removed before the subsequent classification and selection of features to prevent the impact on the classification results.

Chinese word segmentation is an important step to decompose the whole text into basic word-based processing elements. The accuracy of word segmentation affects the accuracy of subsequent classification. At present, Chinese word segmentation algorithms [19] can be divided into four types: (1) Dictionary word segmentation, also known as mechanical word segmentation or matching word segmentation. The algorithm must build a “dictionary” and then compare the text to be segmented with the words in the “dictionary” one by one. If the match is found to be successful, the word is segmented. (2) Understanding lexical segmentation, also known as understanding knowledge-based lexical segmentation. The algorithm is based on syntactic analysis and semantic analysis to eliminate ambiguity, which requires a lot of language knowledge and information. Due to the complexity of Chinese grammar and the immaturity of word segmentation based on knowledge understanding, it has great research prospects. (3) Statistical vocabulary division, as the name suggests, is that from a statistical point of view, words are a stable combination of Chinese characters. The greater the probability of two words appearing at the same time, the more words will be formed. Therefore, the frequency of adjacent words in the text can be counted and used as the basis of word segmentation. Common statistics include word frequency and information. Word frequency is a commonly used weighting technique used in intelligence retrieval and text mining to evaluate the degree of repetition of a word for a document or a set of domain documents in a corpus. Mutual information (mutual information) is a useful information measure in information theory. It can be seen as the amount of information about another random variable contained in a random variable, or that a random variable is due to the fact that another random variable is known. Related word segmentation models include maximum likelihood word segmentation model, maximum entropy word segmentation model, hidden Markov word segmentation model, and directed graph model. Combinatorial dictionary refers to the combination of two or three of the three methods. It can not only overcome the disadvantages and shortcomings of a single word segmentation method but also integrate some algorithms to improve the accuracy and speed of word segmentation. After text preprocessing, the text is represented as structured data that can be processed directly by the computer. The feature selection process includes selecting some feature vectors from the structured data and constructing the best feature subset. (4) Feature selection, also known as independent evaluation method, eliminates features that are useless for

classification by reducing the feature vector dimension of samples and retaining features that are meaningful for classification.

The method of selecting features is usually to establish an evaluation function, evaluate each feature vector and calculate its evaluation score, then sort the features according to the evaluation score, and then select the best  $n$  features as the subset of features. Common methods of feature selection include document frequency (DF), information gain (Ig), chi square ( $\chi^2$ ) statistics, entropy, and expected cross entropy (ECE). Among them, document frequency refers to the number of documents that contain a specific term. In probability theory and information theory, information gain is asymmetric, and it is used to measure the difference between two probability distributions  $P$  and  $Q$ . Information gain describes the difference when using  $Q$  for encoding and then using  $P$  for encoding. The chi square statistic is a measure of the difference between the distribution of the data and the selected expected or hypothetical distribution. Expected cross entropy is used to measure the importance of a word to the whole.

At present, according to the different starting points of text data processing, text classification algorithms are mainly divided into two categories: rule-based classification method, which is used to extract classification rules through a given threshold. Rule-based classification methods appeared earlier. The common methods are decision tree, random forest, association rules, and so on. The statistical-based classification method extracts features to represent the text and transforms the classification into a mathematical problem. Generally speaking, based on machine learning, the known classification samples are trained to obtain the classifier. The typical methods of this method include naive Bayesian algorithm, nearest neighbor  $K$  algorithm, neural network, and SVM algorithm. Among them, the premise of the naive Bayes algorithm is that the attributes are independent of each other. When the data set satisfies this independence assumption, the classification accuracy is high; otherwise, it may be low. In addition, the algorithm has no classification rule output, and the KNN algorithm itself is simple and effective, but in class decision-making, it is only related to a very small number of adjacent samples. Standard SVM is an algorithm designed based on binary classification problems and cannot directly handle multiclassification problems.

The characteristics of text classification task are as follows: (1) in high-dimensional function space, the classification of document set can produce tens of millions of functions. (2) Feature distribution is sparse. When feature words are used to represent documents, the feature dimension is generally very high, and the feature words appearing in documents are only a small part of all feature words. As a result, the frequency of most feature words is 0, so the value of most feature words in the document vector is 0. (3) The characteristics of polysemy and synonyms exist. For example, “Professor” can mean a professional title or a verb that conveys knowledge. (4) The last characteristic is the diversity and complexity of categories and the special diversity of documents, and there are complex links between different

categories, which are intertwined with each other. Their mutual interference will have a certain impact on the performance, computational complexity, and results of the classifier.

### 3. Experimental Design of Multimedia Book Retrieval

**3.1. Image Retrieval Steps.** In the content-based image retrieval system, the retrieval process can be represented in Figure 5.

After extracting the image features from the system database, it is necessary to adopt a method that can evaluate the similarity of feature data. The function of similarity measurement module is to compare the feature data extracted by computer. The closer the two feature data are, the greater the measure of similarity is. The greater the difference between the two feature data, the smaller the measure of similarity. Therefore, in the image search system, distance is mainly used to measure the similarity between the queried image and the queried image. Common similarity search methods in image search are as follows.

**3.1.1. Euclidean Distance.** Euclidean distance is a similar measure commonly used in image search. If the importance of each dimension of the feature vector (usually 2D or 3D) is the same, and there is an orthogonal relationship between each component of the image data, the actual distance between two different feature vectors  $C$  and  $D$  can be measured by distance  $J_1$  or distance  $J_2$ .

The  $J_1$  distance can be expressed by the following formula:

$$J_1 = \sum_{x=1}^M |C_x - D_x|, \quad (18)$$

where  $m$  is the dimension of the eigenvector.  $J_2$  distance calculation formula is as follows:

$$J_2 = \sum_{x=1}^M (C_x - D_x)^2. \quad (19)$$

**3.1.2. Histogram Intersection Method.** Histogram intersection method was first proposed by scholars. In the research of computer vision interaction, they proposed an image histogram matching method, which can effectively reduce the influence of other factors and determine the position of known objects. The formula of histogram intersection distance is as follows:

$$J(M, N) = \sum_{x=1}^M \min(M_x, N_x). \quad (20)$$

During the operation, the number of pixels shared by the  $N$  handles contained in  $P$  and  $Q$  can be calculated from the color histogram. At the same time, the number of pixels in one histogram can be normalized by dividing the calculated

value by the intersection of all histograms. However, its value shall be within the range of  $[0, 1]$ . The normalization formula is as follows:

$$J(P, Q) = \sum_{x=1}^M \min\left(\sum_x P_x, \sum_x Q_x\right). \quad (21)$$

**3.1.3. Quadratic Distance.** Quadratic distance can be measured by introducing color similarity matrix to illustrate the similarity of colors in different images. In the classic IBM system, the variance is used to measure the duplicate of the image when searching and matching the image. The quadratic distance between two color histograms  $m$  and  $N$  can be expressed as follows:

$$N(M, N) = (N - M)^T C (N - M), \quad (22)$$

where  $T$  stands for transpose, which is a mathematical term. Intuitively, all the elements of  $A$  are mirror-reversed around a 45-degree ray starting from the element in the first row and the first column, and the transposition of  $A$  is obtained.

**3.1.4. Mahalanobis Distance.** Markov distance is the distance between similar sample sets obtained by calculating the covariance matrix of the image as follows:

$$J(M, N) = (M - N)_T C^{-1} (M - N), \quad (23)$$

where matrix  $C$  is the covariance matrix of all eigenvectors.

**3.2. Text Retrieval Steps.** There are three important steps in the classification process. The first step is to use the known classification training data set to train the classifier. The second step is to use the known test data set to evaluate and optimize the classification accuracy of the classifier. The last step is to collect the data of unknown classification through the classifier to get the final result of classification. For text classification, text data should also be regarded as structured data that programs can analyze, that is, preprocessing process. The flow chart of text classification system is shown in Figure 6.

**3.3. Retrieval Performance Evaluation Criteria.** At present, there are many search strategies used in image and text search systems. When using different algorithms to describe features or measure similarity, the performance of the system is also different. Therefore, it is necessary to establish a unified evaluation standard to evaluate the research system. In the field of information retrieval, evaluation indexes such as precision, recall, and retrieval time are generally used [20–22].

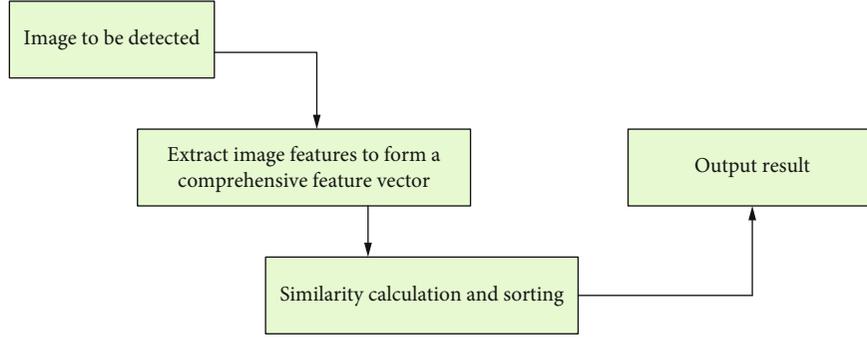


FIGURE 5: Content-based image retrieval system.

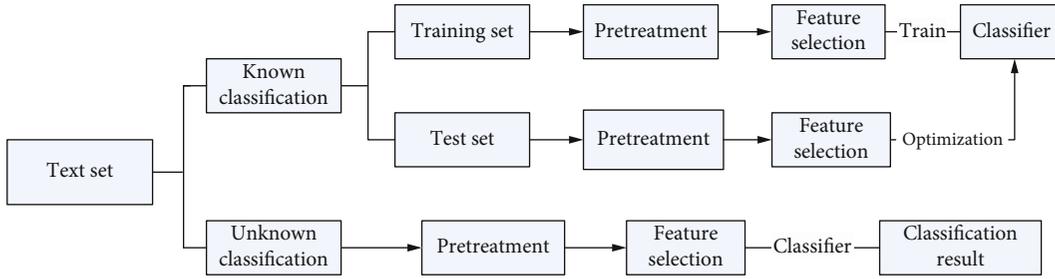


FIGURE 6: Flow chart of text classification system.

TABLE 1: Data selection criteria for different database samples.

	Positive sample	Negative sample	Negative sample draw ratio
Philosophy and religion	201	1243	17%
Political law	410	998	62%
Economics	356	1174	34%
Cultural education	524	816	46%
Natural science	187	1359	12%

TABLE 2: Comparison of selection effects of different text features.

	100	200	500	1000
CHI	62.35%	74.29%	80.12%	85.21%
IG	61.78%	72.37%	75.31%	83.77%
DF	58.46%	64.28%	70.29%	74.28%

TABLE 3: Comparison of selection effects of different image features.

	500	1000	1500	2000	2500
Color characteristics	58.11%	60.35%	61.67%	65.28%	71.46%
Texture feature	51.31%	55.47%	59.84%	62.35%	68.93%
Spatial characteristics	61.35%	65.29%	71.4%	76.37%	79.61%

3.3.1. *Precision and Recall.* Precision and recall are the two most basic evaluation criteria in information retrieval. Their definitions are shown in the following formulas.

$$P = \frac{\text{sum}J}{\text{sum}N}, \quad (24)$$

$$R = \frac{\text{sum}J}{\text{sum}M}, \quad (25)$$

where  $P$  can be expressed as the accuracy of the search and  $R$  can be expressed as the comprehensiveness of the search. The total number of retrieval-related results can be expressed in  $J$ ,  $n$  is the total physical examination result, and  $M$  is the overall related image.

By definition, we can see that the precision is related to the accuracy of search results. Recall is related to the comprehensiveness of search results. Generally, people take the precision  $R$  and recall  $P$  as the  $M$ -axis and  $n$ -axis, respectively, and draw a precision and recall curve by using the points on the coordinate axis, which is called PVR curve. The region around curve  $f(m, n)$  and axis  $m, n$  is called PVR index. The greater the index, the higher the efficiency of the search system.

3.3.2. *Retrieval Time.* In addition, retrieval time is an important index to evaluate the performance of retrieval system. It represents the total time required for the system to collect the characteristics of the picture and check the duplicate after the user submits the picture to be queried. It can be considered as the response time of the search system. In order to ensure the accuracy of search, the faster response

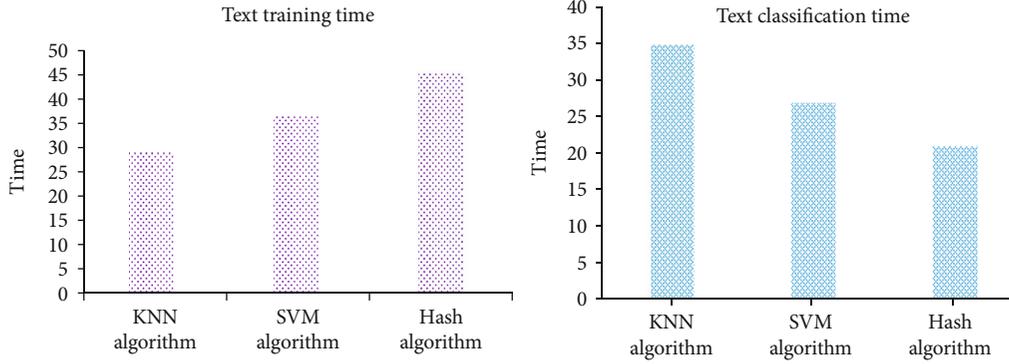


FIGURE 7: Comparison of text training and classification time of each classification algorithm.

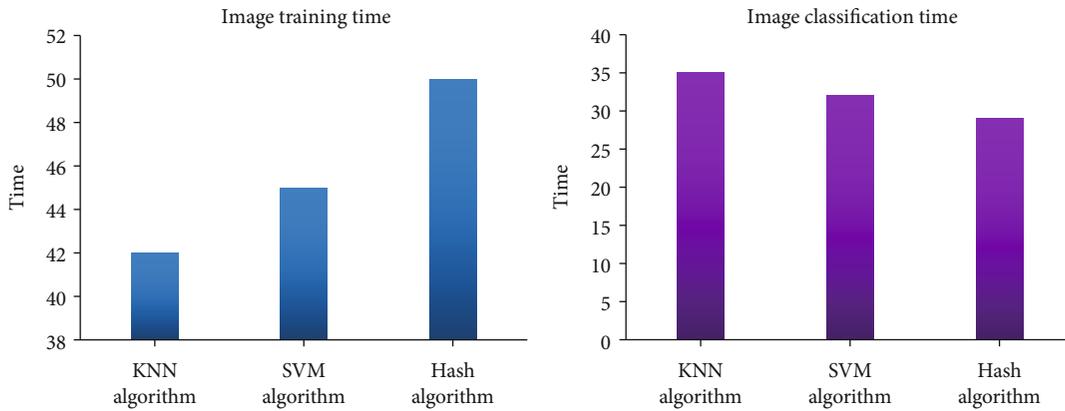


FIGURE 8: Comparison of image training and classification time of each classification algorithm.

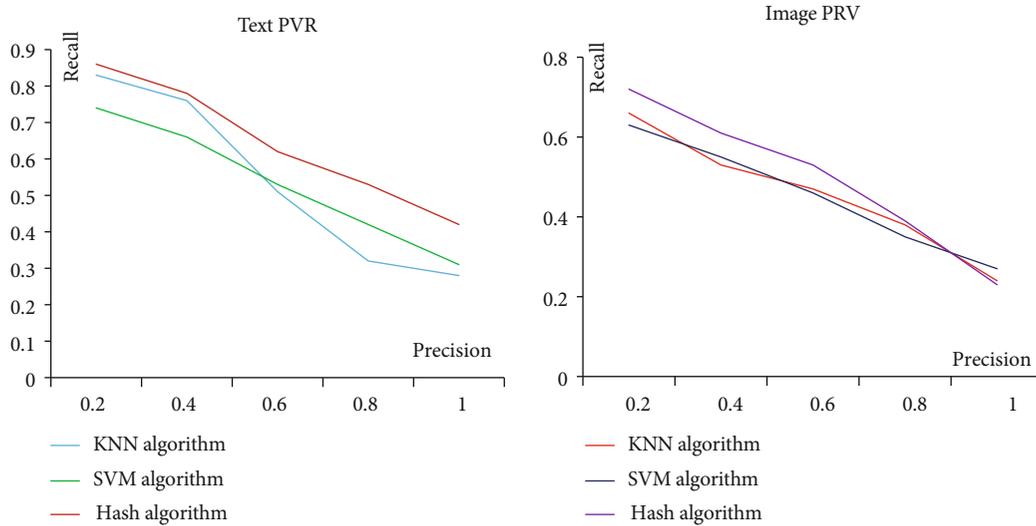


FIGURE 9: Image and text retrieval PVR diagram.

time of the system can enable users to obtain a better search experience.

3.4. *Experimental Design.* The classification experiment selects the cloud media resource book research data set provided by a university library to evaluate the classification effect. Library books include 22 categories, including more

than 15000 book data. The experiment selects five kinds of book data and 6000 book data for the experiment. 60% of the documents are considered as training sets, and the remaining 40% are considered as test sets, as shown in Table 1.

Rocchio algorithm is a very intuitive text classification algorithm. The core idea is to make a standard vector (also

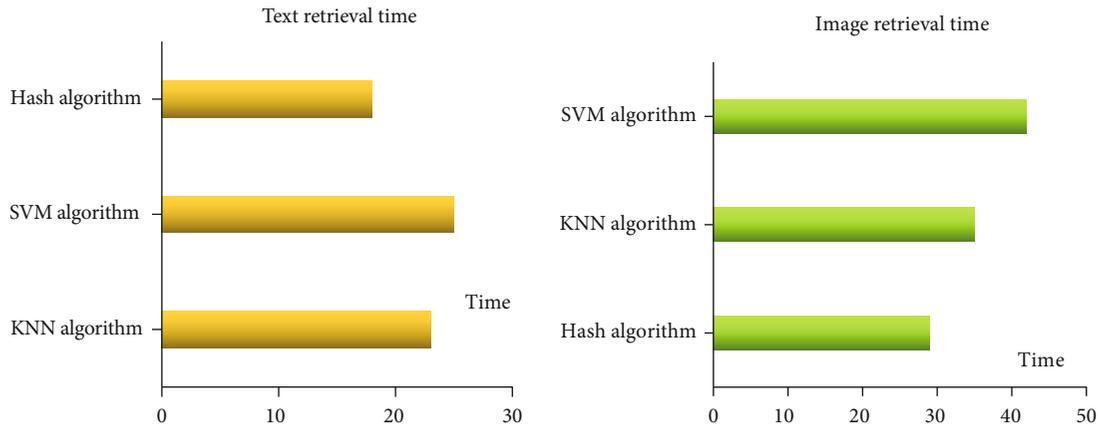


FIGURE 10: Comparison of text and image retrieval time.

called a prototype vector) for each document category and then use the vector of the document to be classified and this standard vector to compare the cosine similarity. The higher the similarity, the more likely it is to belong to this category, and vice versa. Rocchio classification algorithm is selected in the classification experiment. Considering the short length of text samples, the experimental feature dimensions of 100, 200, 500, and 1000 are selected, respectively. According to the experimental design, the results are shown in Table 2.

The results in Table 2 show that the CHI statistical method is the best of the three feature selection methods. Therefore, in the following classification system, the feature selection method of chi square statistics is used to select features.

The experimental dimensions of image features are 500, 1000, 1500, 2000, and 2500. Table 3 shows the specific survey data.

The data in Table 3 shows that spatial features are the best of the three feature selection methods. Therefore, in the following classification system, the method of selecting spatial features extracts features from the image.

According to the test data set, the text extracts relevant features through chi square statistics, and the image extracts relevant features through spatial features. Three common classification algorithms (KNN algorithm, SVM algorithm, and hash algorithm) are selected. Through comparative experiments, select the appropriate classification multimedia book cloud resource algorithm, get the classification search results, and select the optimal classification algorithm.

The results of Figures 7 and 8 show whether text data or images are involved. Compared with KNN and SVM, hash algorithm performs well in classification accuracy. In terms of classification efficiency, the classification training time of hash algorithm is less than KNN and SVM. However, due to the inertia of hash algorithm and the large computational load in the classification process, the classification time is relatively slow. Therefore, comparing the efficiency and accuracy of classification, in the system classification search, select the hash algorithm with better comprehensive classification performance to train the classifier.

As shown in Figure 9, the accuracy and recall of text search are better than complex image search. Compared

with other algorithms, hash algorithm is better than KNN and SVM in all aspects. It shows good search effect and is more suitable for multimedia book cloud resource retrieval.

As shown in Figure 10, the search table time of hash algorithm is significantly less than KNN and SVM, which can better meet the retrieval needs of users. Therefore, hash algorithm can be well used for text or image retrieval and play a greater role in multimedia cloud resource book retrieval.

#### 4. Discussion

This paper studies the classification technology of multimedia cloud resource books. This chapter first introduces the hash algorithm and the relevant principles of cloud resources and multimedia information classification, then introduces the information of relevant text classification, and combs the process of text classification. Combined with the multimedia cloud resource book classification and retrieval studied in this paper, this paper designs experiments to verify the most feature selection method of text and image and the appropriate classification and retrieval method—hash algorithm. These provide a technical basis for the subsequent expansion of multimedia cloud resource book classification and retrieval system. In the course of the experiment, through data comparison, the most suitable image and text feature selection method was obtained. After the optimal feature extraction is completed, the detection parameters of each algorithm are measured by different algorithms, and the comparison results of the exact search curve and the retrieval time fully prove that the Lagrangian algorithm is more suitable for multimedia cloud resource retrieval.

#### 5. Conclusion

Starting with the text and image feature extraction of multimedia cloud resources, this paper analyzes the response effects of different retrieval methods on multimedia resources and finds the most suitable algorithm for multimedia cloud resources—hash algorithm. This paper fully demonstrates the principle of image text feature extraction and

the process of classification, so that readers can better understand the classification process. At the same time, relying on the classification information evaluation standard, this paper quantitatively evaluates the whole experimental results, which can more intuitively show the classification effect of multimedia cloud resource data. But there are also deficiencies: the data collected in this paper is specific data, and the scalability is poor. This paper focuses on image and text retrieval. The follow-up research work can expand the scope of retrieval. This paper is based on the retrieval of small samples, and the carrying capacity of the system is still uncertain. In the future, this paper will continue to improve the retrieval algorithm, continuously expand the scope of retrieval, broaden the field of retrieval, and realize more strengthened retrieval method testing.

### Data Availability

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

### Conflicts of Interest

The author declares no conflicts of interest.

### References

- [1] S. Gardezi, I. Faye, J. Bornot, N. Kamel, and M. Hussain, "Mammogram classification using dynamic time warping," *Multimedia Tools & Applications*, vol. 77, no. 3, pp. 3941–3962, 2018.
- [2] J. Zhang and J. Jiang, "Decomposition-based tensor learning regression for improved classification of multimedia," *Journal of Visual Communication & Image Representation*, vol. 41, pp. 260–271, 2016.
- [3] S. Biasotti, A. Cerri, M. Aono et al., "Retrieval and classification methods for textured 3D models: a comparative study," *The Visual Computer*, vol. 32, no. 2, pp. 217–241, 2016.
- [4] F. Narducci, C. Musto, M. D. Gemmis, P. Lops, and G. Semeraro, "TV-program retrieval and classification," *Information Systems Frontiers*, vol. 20, no. 6, pp. 1157–1171, 2018.
- [5] T. Adhikary, A. K. Das, M. A. Razaque, M. Alrubaian, M. M. Hassan, and A. Alamri, "Quality of service aware cloud resource provisioning for social multimedia services and applications," *Multimedia Tools and Applications*, vol. 76, no. 12, pp. 14485–14509, 2017.
- [6] Y. F. Liu, J. M. Guo, and L. An, "Multimedia classification using bipolar relation graphs," *IEEE Transactions on Multimedia*, vol. 19, no. 8, pp. 1860–1869, 2017.
- [7] S. Singaravelan, A. S. Balan, and D. Murugan, "WEB based document retrieval using advanced CBRS," *Review of Innovation and Competitiveness*, vol. 3, no. 4, pp. 103–116, 2017.
- [8] F. Al-Akashi and D. Inkpen, "Real-time web search framework for performing efficient retrieval of data," *Journal of Information and Organizational Sciences*, vol. 45, no. 1, pp. 287–308, 2021.
- [9] A. Khamparia, G. Saini, B. Pandey, S. Tiwari, D. Gupta, and A. Khanna, "KDSAE: chronic kidney disease classification with multimedia data learning using deep stacked autoencoder network," *Multimedia Tools and Applications*, vol. 79, no. 47–48, pp. 35425–35440, 2020.
- [10] X. Zhu, J. Yuan, Y. Xiao, Y. Zheng, and Z. Qin, "Stroke classification for sketch segmentation by fine-tuning a developmental VGGNet16," *Multimedia Tools and Applications*, vol. 79, no. 45–46, pp. 33891–33906, 2020.
- [11] H. Kim, Y. Kim, and H. Chang, "Information security research classification for future multimedia environment," *Multimedia Tools and Applications*, vol. 75, no. 22, pp. 14795–14806, 2016.
- [12] M. Amiri and L. Mohammad-Khanli, "Survey on prediction models of applications for resources provisioning in cloud," *Journal of Network and Computer Applications*, vol. 82, pp. 93–113, 2017.
- [13] M. T. Krieger, O. Torreno, O. Trelles, and D. Kranzlmüller, "Building an open source cloud environment with auto-scaling resources for executing bioinformatics and biomedical workflows," *Future Generation Computer Systems*, vol. 67, pp. 329–340, 2017.
- [14] W. Tao and D. Chang, "News text classification based on an improved convolutional neural network," *Tehnicki vjesnik-Technical Gazette*, vol. 26, no. 5, pp. 1400–1409, 2019.
- [15] Z. A. Baig, S. M. Sait, and F. Binbeshr, "Controlled access to cloud resources for mitigating Economic Denial of Sustainability (EDoS) attacks," *Computer Networks*, vol. 97, pp. 31–47, 2016.
- [16] H. Hallawi, J. Mehnen, and H. He, "Multi-capacity combinatorial ordering GA in application to cloud resources allocation and efficient virtual machines consolidation," *Future Generation Computer Systems*, vol. 69, pp. 1–10, 2017.
- [17] X. Yang, M. Guo, Q. Lyu, and M. Ma, "Detection and classification of damaged wheat kernels based on progressive neural architecture search," *Biosystems Engineering*, vol. 208, no. 1, pp. 176–185, 2021.
- [18] F. Isiaka, K. S. Mwitondi, and A. M. Ibrahim, "Detection of natural structures and classification of HCI-HPR data using robust forward search algorithm," *International Journal of Intelligent Computing and Cybernetics*, vol. 9, no. 1, pp. 23–41, 2016.
- [19] W. Wang, "Cloud manufacturing resources semantic description and service matching strategy," *Chongqing Daxue Xuebao/Journal of Chongqing University*, vol. 40, no. 5, pp. 1–6, 2017.
- [20] J. Y. Yoon and S. Joung, "A big data based cosmetic recommendation algorithm," *Journal of System and Management Sciences*, vol. 10, no. 2, pp. 40–52, 2020.
- [21] I. Belosevic and M. Ivic, "Variable neighborhood search for multistage train classification at strategic planning level," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 3, pp. 220–242, 2018.
- [22] P. Lisanna, H. Layla, P. Damiano, M. A. Andrade-Navarro, A. V. Kajava, and S. Tosatto, "RepeatsDB 2.0: improved annotation, classification, search and visualization of repeat protein structures," *Nucleic Acids Research*, vol. 45, no. 6, pp. 3613–3613, 2017.