Research on Online Foreign Language Teaching Based on CBIR Algorithm of Internet Deep Information Retrieval

Hou Jun
Xi`an International Studies University
Xi`an , 710028, Shaanxi, China

Abstract:This paper proposes a novel content-based image retrieval CBIR algorithm model, and also proposes a novel clustering retrieval method suitable for image data, which improves the clustering accuracy by automatically updating the weights of cluster features, thereby improving the retrieval effect. It aims to build a reasonable online teaching model, effectively guide foreign language online teaching in colleges and universities, provide reference for mixed teaching research, and promote students to improve their comprehensive foreign language skills and optimize their learning outcomes. The changes in the grades of the experimental class were also analyzed separately, and the results of the experiment were supplemented by post-study questionnaires and interviews.

Keywords: Online Foreign Language, Language Teaching, CBIR Algorithm, Internet Deep Information Retrieval

1. INTRODUCTION

The retrieval of the traditional CBIR system is to return the results according to the similarity of the underlying features. Often [1], there are some images in the front that although the underlying visual features are very similar, their semantic content and the retrieved images are different. match. Image retrieval [2] is developed from text information retrieval technology. Due to the differences between the characteristics of images and texts, many successful indexing techniques in text retrieval [3] technology cannot be applied to image retrieval. in-depth research. The research of image retrieval is divided into two categories: Text Based Image Retrieval (TBIR) and Content-Based Image Retrieval (CBIR) [4].

Text-based image retrieval mainly uses text annotation to describe the content of the image as a whole [5], that is, the keywords of the image. Education informatization promotes education development, and fair development of artificial intelligence will gradually change the spatial [6] limitations of high-quality educational resources, making it possible to share high-quality educational resources across regions, and promoting the universal [7] sharing of high-quality resources. Most image search engine websites, such as Google, Baidu, etc. generally use this method [8].

However, there are two major problems in this image retrieval technology: First, it is time-consuming and labor-intensive to manually [9] label images, especially in the face of massive image databases, the workload of manual labeling is huge. In this model, an improved clustering algorithm (Improved Cluster for short IC) [10] is used, and the semantic retrieval is combined with the retrieval of the underlying visual features, which improves the accuracy of image retrieval. Then, document analysis technology is used to realize image retrieval [11], but because of the vector quantization method used in defining key image blocks, the independence between key sub-blocks cannot be guaranteed. And the algorithm for generating the codebook is complex and time-consuming, thus affecting the final retrieval effect [12].

The early studies of these CBIR systems mainly focus on retrieval by extracting rich visual features (eg, color, texture, and shape) of images [13]. Following these earliest studies, there are some other studies such as Ju Han and Kai-KuangMa [14] who proposed a method to blur color histograms. Artificial intelligence technology is widely used in English teaching, and has been able to undertake some of the teaching functions of teachers [15]. How teachers work with intelligent teachers to improve English teaching in junior high schools. As the main body of learning in colleges and universities, students can make full [16] use of network information and new media technology to strengthen online foreign language learning, carry out online cultural exchanges, optimize learning results, and promote the development of comprehensive abilities [17].

The application of example technology in the field of image retrieval has been deeply studied, and the multi-example technology is applied in the fields of image annotation [18], image retrieval and face recognition to realize image retrieval based on object semantics. The significance of the research is mainly. The development process of the flipped classroom can be divided into three stages [19]: from the budding stage to the conception, to the slow development stage and the rapid development stage (Zhang Wei, 2017). Therefore [20], the author will use time as the main line to sort out the foreign flipped classroom in detail. literature and review. The CBIR model proposed in this paper covers image retrieval and automatic semantic annotation.

Divided into two different cases, two different retrieval methods based on visual content and based on semantic keywords are carried out. Its basic structure is shown in Figure 1. We must also pay attention to the internal laws of foreign language teaching [21], clarify the relationship between language and culture, and teach language from a cultural perspective. Starting from the relationship between language and culture, this paper explains the importance of cultural introduction in [22] online foreign language teaching, combined with case analysis and the current situation of online foreign language teaching during the period of delayed start of school due to epidemic prevention [23].

Based on this, this paper proposes an image retrieval algorithm based on text retrieval technology. Combined with

the visual characteristics of the human eye, image sub-features that are meaningful to characterize the image content are constructed [24]. Because traditional low-level features are not only relatively cumbersome to extract, but also relatively complicated to follow-up processing of features, which brings great inconvenience to content-based image retrieval. However, the current popular hashing method has defects in data storage and calculation, so this paper proposes a method based on CNN and graph theory for image retrieval [25].

2. THE PROPOSED METHODOLOGY

2.1 The Internet In-Depth Information Retrieval

Users need to submit semantic keywords and example images for querying. According to the example images, the database images without the keyword annotation are searched based on the visual feature sub-module, and the database images that have been marked with the keyword are semantic-based Keyword Retrieval. Block encoding is an effective and fast lossy digital image compression technique. It is easier to implement than vector encoding, and the algorithm is simple and preserves the edge information of the image. Since the visual information sensed by the sensory cells of the human cerebral cortex is directional, and the convolutional neural network can extract more abstract image features, it is widely used in the CBIR system.

The basic principle of convolutional neural network is: given an image library, given a convolutional network model, optimize a cost function to train optimal weights, and extract image features through optimal weights. It is unrealistic to manually annotate the massive image data on the Internet. With the continuous maturity of information retrieval technology, automatic collection and indexing of web page information, as an important part of search engines, has been deeply studied and widely used in text search engines. Another case is the semantic retrieval module (Fig. 1), which targets already semantically annotated images. The retrieval in this case is relatively simple, and the results can be returned directly according to the confidence of a certain semantic order, which is similar to the ordinary keyword-based retrieval. Suppose I is a grayscale image of size M×N.

First, divide J into non-overlapping sub-blocks of m×m size. For each sub-block, calculate the average grayscale value and average color difference of the pixels in the block. Then according to the idea of block coding, VGG-net can be said to be a deepened version of Alex-net at the network level. It includes 5 network structures, called A, B, C, D, E, and each network structure has different The number of network layers, and the number of network layers from A to E gradually increases, namely: 11, 13, 16, 16, 19 layers. VGG-net and Alex-net are similar in overall structure. Vector Space Model (VSM) is widely used in text mining. Based on this model, the potential concepts of text and the relationship between concepts can be discovered. Discover hidden knowledge. Usually calculated by the TF/IDF formula.

2.2 The Information Retrieval CBIR Algorithm

Therefore, in the algorithm, a threshold is set. When the average color difference of an image block is less than this threshold, the block is regarded as a uniform block. The content-based image retrieval is mainly through the user inputting a picture, and then from the image library The

process of finding images similar to the input image and recommending them to users. The learning system builds a model by learning on packages of already labeled categories, hoping to predict the concept labels of unknown packages as correctly as possible. Due to its unique nature, multi-instance learning has attracted the attention of scholars and is considered to be the fourth learning framework alongside supervised learning, unsupervised learning and reinforcement learning. Content-based image retrieval firstly uses the corresponding feature extraction method to extract the features of the images in the image library, and establishes the index between the image library and the corresponding features.

For a given query image, take the cluster center image of each cluster as the core, and calculate the distance between each image and the cluster center according to the image distance formula with the feature weight adjusted. The images are clustered based on the minimum distance criterion. Based on the above evaluation criteria, two different sets of comparative experiments are carried out.

2.3 The Research On Online Foreign Language Teaching

First, experiments are conducted on the effect of different block methods on the retrieval performance of the algorithm when extracting sub-features. When extracting sub-features, the block method, that is, the selection of sub-block size, directly affects the effectiveness of image content description.

Teachers can record micro-lectures about Baby boomers, Generation X, and Millennials before class, and publish them to students through the Internet. Students can express their views on the groups these words represent and the origin of these names in the Q&A discussion forum. In the experiment, when using formula (2) to measure the similarity, we take one mouth and one 0.5, so that the similarity of the number of features and the similarity of features are equally important.

Since VGG-net [11] has achieved good results in the classification of ImageNet in 2012, this paper uses the VGG-net model to extract image features, but the feature dimension extracted by VGG-net is higher, in order to better Using the feature space, the feature is firstly dimensionally reduced by kernel principal component analysis. Although the school does not open, it does not stop, and it is possible to teach courses at a distance. The advancement of information technology and network platforms provides a great possibility for the education circle of our country to face the education dilemma of isolation and home. The multi-modal online output of foreign language teaching can break through the time limit of traditional foreign language classes. Teachers can send homework reminders to students and monitor students' learning progress and knowledge mastery at any time.

Through the online interactive platform, a multi-modal online discussion is conducted with students on a certain issue, and students can submit texts to teachers individually. However, the scores of the experimental class were higher than those of the control class, indicating that the English writing score of the students in the experimental class after the study was significantly higher than that of the students in the control class after the study, indicating that flipped classroom teaching can significantly improve the students' English writing scores. In order to test the complexity of the algorithm in this paper, the algorithm in this paper is compared with the other two methods in terms of the time complexity of feature

extraction and the time complexity of image retrieval under the same hardware and software environment.

3. CONCLUSIONS

Big data has brought new ways for colleges and universities to innovate education management models. Colleges and universities should arrange all-round campus perception terminals on the basis of campus informatization construction. Sort out the correlation between different behaviors and draw students' "student portraits" in school. , clearly describe the situation of students in school, and analyze the relationship between students' learning situation, life status and also psychological dynamics. Clean, integrate, mine and apply massive, heterogeneous and multi-dimensional campus data, and extract potential, valuable, and potentially valuable information from it.

4. ACKNOWLEDGEMENT

5. REFERENCES

[1]Meng Fanjie, Song Miao, Shan Dalong, et al. Web Image Retrieval Method Based on Semantic Propagation and Hybrid Multi-Instance Learning: CN106202256B[P]. 2019.

[2] Zhen Junjie, Ying Ziluo, Zhao Yihong, et al. Image retrieval based on unsupervised retraining convolutional neural network [J]. Modern Computer, 2019(14):5.

[3] Du Anyu. Research on multi-feature image retrieval algorithm based on privacy protection [D]. Xinjiang University, 2019.

[4] Sun Qiping. Research on Image Retrieval Based on Deep Learning [J]. Journal of Jingdezhen University, 2018, 33(3):4.

[5] Zhou Xianchun. Design of Maximum Frequent Subgraph Similarity Matching System Based on Fuzzy Graph Neural Network [J]. Modern Electronic Technology, 2021.

[6] Zhang He. Image retrieval algorithm based on color and texture [J]. Radio and Television Information, 2020(2):4.

[7] Miao Zhuang. Research on the application algorithm of entropy in the field of image retrieval [D]. Jilin University, 2019.

[8] Zhao Shan, Tang Yongli. Research on CBIR Algorithm Based on Text Retrieval Technology [J]. Journal of Optics, 2009(10):5.

[9] Ge Yanbin. Research on accounting informatization teaching in secondary vocational schools based on "Internet +" [J]. Century Star - Exchange Edition, 2021(26):2.

[10] Fu Mingduan. \"Internet +\" background English teaching on the cultivation of students' critical thinking ability [J]. New Education Times Electronic Journal (Teacher Edition), 2018, 000(038):249-250.

[11] Zhang Zhiqiang. A new CBIR system research based on improved clustering retrieval algorithm [J]. Computer Science, 2008.

[12] Luo Weifang. Research on content-based image retrieval technology [D]. South China University of Technology.

[13] Hong Xiaobing. Research on English teaching innovation based on the "Internet +" environment [J]. Campus English, 2018(15):2.

[14] Li Xiang. Research on College English Teaching Based on Online Classroom Teaching Mode [J]. Campus English, 2020(36):2.

[15] Huang Bing. Strategies for cultivating in-depth learning ability of English majors in the era of "Internet +" [J]. Scientific Consulting, 2021.

[16] Gui Chuanghua. Web Image Retrieval Based on Semantic Analysis [J]. 2010.

[17] Wenquan Che, Peng Hong, Li Qiong. Research on key technologies of content-based image retrieval [J]. Microcomputer Information, 2007, 023(003):278-280.

[18] Ma Chao. Research on CBIR system based on NSCT [D]. Lanzhou University, 2007.

[19] Li Qian. Research on deep learning strategies of college English in the era of "Internet +" [J]. 2021(2017-12):386-386.

[20] Wei Ruiping. Research on the strategy of English informatization teaching in junior high school under the background of "Internet +" [J]. Caizhi, 2020(1):1.

[21] Lu Li. Exploration and practice of college English grammar teaching based on "Internet +" [J]. Writer's World, 2020(6):2.

[22] Yu Jie. Research on English teaching in colleges and universities under the background of "Internet +" - Comment on "Internet + English Teaching" [J]. Contemporary Education Science, 2020(11):1.

[23] Wu Xiaoli. Research on higher vocational English smart teaching based on "Internet +" [J]. Journal of Jiamusi Vocational College, 2018(11):2.

[24] Wang Zhen. Current Situation and Promotion Strategy of Information Literacy of Foreign Language Teachers in Higher Vocational Colleges under the Background of "Internet +"——Based on a Survey in Guangzhou [J]. Guangzhou Vocational Education Forum, 2018.

[25] Wang Jiying. Research on Content-Based Image Retrieval Algorithms [D]. University of Science and Technology of China, 2012.