Academic Journal of Computing & Information Science
ISSN 2616-5775 Vol. 6, Issue 11: 83-87, DOI: 10.25236/AJCIS.2023.061111

Application Analysis of Target Image Retrieval Technology in Computer Vision

Tebo Dai

School of Computer Science and Engineering, Hunan University of Science and Technology, Xiangtan, China
tb_dai@163.com

Abstract: *Computer vision is a field of study that explores how computers can understand and perceive their surroundings and objects by processing visual information, such as images or videos. Among its applications, image retrieval is a crucial direction that allows users to quickly find similar images or related information by inputting a single image. This article takes image retrieval as the starting point, defining its basic principles and delving into the practical applications of target image retrieval technology in computer vision. Its aim is to classify and cluster images, enabling faster and more accurate searches for desired images.*

Keywords: *Computer vision, image retrieval, attention models, application analysis*

1. Introduction

Image retrieval technology is a content-based search technique that involves comparing a target image with a large database of images to find the most similar ones. In the field of security surveillance, image retrieval technology can assist law enforcement in quickly and accurately identifying suspects or targets. In medical image analysis, it can help doctors swiftly and accurately diagnose medical conditions. In the realm of autonomous driving, image retrieval technology aids vehicles in precisely recognizing road conditions and obstacles. Additionally, in video content management, this technology enables users to rapidly and accurately locate the video content they need.

2. Basic Overview of Image Retrieval Image retrieval

Primarily involves converting images into digitized feature vectors and then determining the relevance between images by comparing the similarity of these feature vectors [1]. Common image features include color features, texture features, and shape features, among others. These features can be extracted and described using various algorithms and techniques, such as color histograms, gray-level co-occurrence matrices, and edge detection. The process of image retrieval typically consists of two stages: indexing and querying. During the indexing stage, the system extracts and describes features for all images in the image library and stores them in a database. Consequently, when a user inputs a query image, the system can find similar images by calculating the similarity between the query image's feature vector and the feature vectors of all images in the database. As shown in Figure 1 and Figure 2, two similar images can be used as an indexed image and a query image, respectively, in the indexing and querying stages. Image retrieval finds a wide range of applications. In the medical field, it helps doctors quickly identify cases similar to a patient's condition, thereby assisting in diagnosis and treatment decisions. In the security sector, image retrieval is used for facial recognition and object detection, enhancing recognition capabilities and crime prevention. In the entertainment industry, image retrieval is employed for image search and recommendations, helping users discover and access content of interest.

ISSN 2616-5775 Vol. 6, Issue 11: 83-87, DOI: 10.25236/AJCIS.2023.061111

Figure 1: Original image of the car

Figure 2: Car Query Graph

3. Computer Vision Attention Models and the Necessity of Consistency in Target Image Retrieval Schemes

3.1. Analysis of Attention Models

In computer vision, attention models contribute to improving the performance of image retrieval. Image retrieval involves finding images in an image database that are similar to a user's query input. Traditional image retrieval methods typically rely on global image features such as color histograms and texture features. However, these methods often overlook crucial local information within images, leading to inaccurate retrieval results [2]. To address this issue, consistency in target image retrieval schemes introduces attention models. This scheme generates multiple attention maps within an image, restricting attention to regions relevant to the query. These attention maps can be generated using various techniques, including saliency-based methods and deep learning-based methods [3]. Methods for generating attention maps can be broadly categorized into two types: bottom-up and top-down. Bottom-up methods start by decomposing the image into a series of basic visual elements or features, such as color patches, edges, and corners. Then, by analyzing the combination and distribution of these basic elements, they calculate the contribution or saliency of each pixel to the overall image. This calculation resembles a bottom-up construction process, starting from basic visual elements and gradually building up a more complex image understanding. In contrast, top-down methods work in the opposite direction. They begin with high-level semantic information, such as the results of object detection or image segmentation, and then refine this information gradually to form attention maps. This approach first identifies high-level features or objects in the image, such as faces or vehicles, and then computes the importance of each pixel based on the position and distribution of these high-level features. This calculation follows a top-down parsing process, starting from high-level semantic information and gradually breaking it down into more specific pixel-level attention.

3.2. Necessity of Retrieval Schemes

Firstly, it can enhance the accuracy of image retrieval. Traditional image retrieval methods often rely solely on keywords or tags to retrieve images and may not accurately understand the content of the images. In contrast, computer vision attention models and consistent target image retrieval schemes can

ISSN 2616-5775 Vol. 6, Issue 11: 83-87, DOI: 10.25236/AJCIS.2023.061111

accurately identify images that are most similar to the target image by analyzing visual features and structural information within the images.

Secondly, it can improve the efficiency of image retrieval. Since computer vision attention models help us better understand the content of images, they can reduce the search space and enhance retrieval efficiency.

Lastly, it can expand the scope of image retrieval applications. Traditional image retrieval methods are often limited to keyword or tag-based searches and have a limited range of applications. However, computer vision attention models and consistent target image retrieval schemes, by analyzing the content and structural information of images, can be applied to a wider range of image retrieval scenarios.

4. Target Image Retrieval Techniques in Computer Vision

4.1. Feature Extraction

Feature extraction is a critical step in target image retrieval, aiming to extract representative feature vectors from images. Various methods for feature extraction exist, including edge-based methods, localization-based methods, and dimensionality reduction-based methods. Edge detection methods extract features by identifying edges in images, while localization methods focus on extracting features based on local characteristics of the image. Dimensionality reduction methods reduce the dimensionality of data to make feature extraction more efficient and accurate.

4.2. Similarity Comparison

Similarity comparison involves comparing the extracted image features with features from a database of images to identify images with high similarity. Common similarity comparison methods include position and shape-based similarity comparison, and feature matching-based similarity comparison. These methods effectively measure the similarity between images, providing the foundation for subsequent retrieval processes.

4.3. Deep Learning

Deep learning plays a crucial role in target image retrieval. Key concepts include Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Recursive Neural Networks (RNNs), among others. In target image retrieval, deep learning can be used for tasks such as feature extraction and similarity comparison. By training deep neural networks, high-level feature information can be learned from images, enhancing the accuracy and efficiency of image retrieval [4].

4.4. Multimodal Fusion

Multimodal fusion involves combining information or images from different modalities to obtain a richer feature representation. These different modalities may originate from different sensors, time points, perspectives, and more. Multimodal fusion methods include offline training and online training. Offline training involves pre-training and learning from various modalities before merging them, while online training directly combines multiple modalities for training and learning.

4.5. Semantic Understanding

In target image retrieval, semantic understanding helps the system better comprehend the user's query intent, thereby improving retrieval accuracy and efficiency. Common semantic understanding methods include bag-of-words models, speech recognition, and sentence-level models. The bag-of-words model is a text-based image classification method, speech recognition translates speech into text, and sentence-level models provide higher-level semantic descriptions of images.

4.6. Global Search

Global search refers to the process of searching for targets within a large image database. To enhance search efficiency, global search requires the use of efficient indexing and search algorithms.

ISSN 2616-5775 Vol. 6, Issue 11: 83-87, DOI: 10.25236/AJCIS.2023.061111

Common global search algorithms include keyword-based search and language model-based search. Keyword-based search involves searching for relevant images by inputting keywords, while language model-based search uses natural language descriptions to search for relevant images.

4.7. Region-based Search

In region-based search, the target image is first segmented into regions, and similarity comparison and retrieval are then performed based on the segmented region information. Common region-based search methods include edge and texture-based region segmentation and clustering-based region segmentation. These methods segment images into different regions based on feature information, allowing for the rapid identification of regions similar to the target image.

4.8. Multi-Task Learning

Common multi-task learning includes super-resolution and optical flow computation, among others. Super-resolution techniques enhance the resolution of images, resulting in clearer and more accurate retrieval results. Optical flow computation can be used in tasks such as target detection and tracking, providing additional contextual information for image retrieval.

5. Applications of Target Image Retrieval Technology in Computer Vision

5.1. Experimental Methods

In target image retrieval technology, a common approach is to extract feature vectors from images and then perform image matching by calculating the similarity between feature vectors. Common feature extraction methods include local feature descriptors (such as SIFT, SURF), global feature descriptors (such as color histograms, color moments), and deep learning methods (such as Convolutional Neural Networks). To evaluate the performance of target image retrieval technology, the following experimental methods can be employed:

(1) Dataset Preparation: Prepare a database containing multiple target images. The images in the database should exhibit diversity, including different scenes, angles, and lighting conditions.

(2) Feature Extraction: For each target image, use an appropriate feature extraction method to extract its feature vector. Multiple feature extraction methods can be selected for comparison.

(3) Similarity Calculation: For a query image, use the same feature extraction method to extract its feature vector. Then, calculate the similarity between the query image and each image in the database. Common similarity calculation methods include Euclidean distance, cosine similarity, and others.

(4) Evaluate the performance of target image retrieval technology by comparing it with the true labels in the database. Common evaluation metrics include accuracy, recall, F1 score, etc. Through these experimental methods, you can assess the performance of target image retrieval technology on different datasets and select the most suitable feature extraction and similarity calculation methods. Furthermore, you can further optimize target image retrieval technology, such as incorporating deep learning methods for feature extraction and similarity calculation, to improve retrieval accuracy and efficiency.

5.2. Application Effect Evaluation

Target image retrieval technology has widespread applications in the security field. For example, security systems can recognize and retrieve targets through surveillance camera images. When suspicious individuals enter the monitoring range, the system can quickly retrieve images similar to them, enabling timely security measures.

In the field of medical image diagnosis, target image retrieval technology also shows excellent results. Doctors can input a patient's medical images as query images, and the system can quickly retrieve similar cases for reference. This allows doctors to make more accurate diagnoses and provide appropriate treatment plans.

Moreover, target image retrieval technology can be applied in the realm of product recommendations. For instance, on e-commerce platforms, users can upload an image of a product they

ISSN 2616-5775 Vol. 6, Issue 11: 83-87, DOI: 10.25236/AJCIS.2023.061111

like, and the system can retrieve similar product images, recommending them to the user.

6. Conclusions

In conclusion, in-depth analysis and the application of target image retrieval technology based on computer vision hold significant significance. Moreover, in the era of information explosion, web-based image retrieval deserves comprehensive research, with human-computer interaction and related feedback emerging as hot topics in image retrieval research. Multi-modal retrieval methods can fundamentally address the challenges faced in target retrieval, enhancing its practical value.

References

[1] Zhao Zihan. (2021). Image Retrieval Technology Based on Multi-Feature Fusion. Electronic Technology and Software Engineering, (14), 125-126.

[2] Yu Laixin. (2019). Research on Image Retrieval Methods Based on Visual Attention Models. Published Doctoral dissertation, Dalian University of Technology, Dalian, Liaoning.

[3] Liu Qing. (2021). Research on Image Retrieval Methods Based on Deep Learning. Published Master's thesis, Qingdao University of Science and Technology, Qingdao, Shandong..

[4] Huang Jiesheng. (2022). Research on Image Retrieval Algorithms in Computer Vision Based on Deep Learning. Information Technology and Informatization, (09), 181-184.