

An Improved Image Descriptor for Image Classification and CBIR Applications



M. Venkata Dasu, M. Guru Sohitha, I. Giri Vardhan, B. Bindu,
and T. Abhilash

Abstract Recent technological advancements have significantly expanded multimedia complexity, opening up new research fields that rely on similar multimedia material retrieval for image classification and CBIR. Using image features from those images, such as color, composition, structure, and spatial configuration, rather than manually annotated textual keywords, content-based image retrieval (CBIR) searches for images in a massive dataset that are comparable to a given query image. In this methodology, a histogram of oriented gradients (HOG), Color Histogram along with HSV are used for extracting shape and color. The Gabor Wavelet Filter has additionally been applied to object and texture recognition. And an improved convolutional neural network (CNN) is used for similarity matching. In this project, the experiments run on image database having 19 categories and each category contains one hundred images.

Keywords Convolutional neural network (CNN) · Image classification · Content-based image retrieval (CBIR) · Image descriptor · Histogram of oriented gradients (HOG) · Gabor wavelet

1 Introduction

In 1992, Kato coined the phrase “CBIR” for the first time. Other names for content-based image retrieval (CBIR) include search by image content and content-based visual information retrieval (CBVIR). A CBIR’s main objective is to use the visual content of the photos to discover relevant images within a vast image library. Moreover, CBIR decreases the similarspaceproblem, which enhances the efficiency of

M. Venkata Dasu (✉) · M. Guru Sohitha · I. Giri Vardhan · B. Bindu · T. Abhilash
Department of ECE, Annamacharya Institute of Technology and Sciences, Rajampet, Andhra Pradesh, India
e-mail: dassmarri@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024
A. Kumar and S. Mozar (eds.), *Proceedings of the 6th International Conference on Communications and Cyber Physical Engineering*, Lecture Notes in Electrical Engineering 1096, https://doi.org/10.1007/978-981-99-7137-4_67

689

image retrieval. Previously, searching an image database relied on human annotation, in which each image in the database was given a set of keywords to indicate its semantic meaning [1].

This type of image retrieval is called as Text Based Image Retrieval. After that, images are indexed using all the keywords. This TBIR approach has many drawbacks, including the growing size of the image collection and the difficulty of manually annotating each image. It is individual to annotate an image based on someone's perception. Content-based image retrieval was proposed to overcome the limitations of text-based image retrieval. CBIR has since been utilized as an alternative for text-based image retrieval. The user interface is crucial since it serves as the main interaction tool.

In general, the term "Content-based image retrieval" (CBIR) is used to describe a retrieval method that uses contents named as low-level features or visual elements to retrieve and get images from a huge image library in response to user requests that are given in the form of query images. The most frequent components of visual information include colors, textures, forms, and the spatial configuration of a region of interest. query image, it can identify pertinent pictures that are like the query image. Due to the close visual resemblance of the visual attributes of these photos, extracting pictures from a database using these features is difficult.

Generally, CBIR is carried out in two steps. Searching and indexing are two of them. When an image is indexed, its features or contents are extracted and stored in feature databases as feature vectors. This applies to both the image being queried and the images in the image database. Feature extraction is the name of this procedure. In CBIR, features can be extracted using a number of different techniques. The characteristics that may be retrieved using CBIR include color, structure, and form.

There are numerous extraction techniques for each of these distinguishing traits. Color space, color histograms, and color moments are frequently used in color extraction techniques. In order to extract the database entries that most closely resemble the query image, a user-constructed input image feature vector is evaluated to every feature vector in the database for resemblance.

The CBIR technology has a wide range of potential applications [2]. They include defense, the military, and intellectual property. Architectural and engineering design, clothing and interior design, journalism and advertising, medical diagnosis, geographic information and remote sensing systems, cultural heritage, education and training, at-home entertainment, and web searching are just a few examples of the various fields that fall under this category.

1.1 Features in CBIR

In a picture, features are recognizable patterns that signify the most important information. Features describe and organize the content of a picture. These are known as the image's properties or characteristics. To achieve a high classification performance in image processing, characteristics are frequently used for search, retrieval,

and storage. Images may be used to extract a variety of features, such as color, texture, and form.

1.1.1 Colour Features

Color is one of the most widely utilized low-level visual elements of CBIR. As the human visual system can discriminate between various sights depending on their hues, color is one of the most significant factors that researchers regularly employ in their work. Color properties are computed using color spaces. The colors that are most often used in the content-based image retrieval domain include HSV, YCbCr, and RGB, as well as LAB. These color spaces are described using color moments, color correlograms, color histograms, dominant color descriptors, color co-occurrence matrices, and several more descriptors [3].

1.1.2 Textural Features

Textures are observable patterns that cannot be represented by just one intensity or hue. As the texture is a common feature in actual pictures, it is an important part of computer vision (CV), but it is also commonly used in feature extraction and pattern detection. One of texture-based image retrieval's main drawbacks is how sophisticated the analysis is. Another is how sensitive it is to noise. Several methods, including the Gabor filter, Markov random field, wavelet transform, steerable pyramid decomposition, gray-level co-occurrence matrix (GLCM), and edge histogram descriptor, are employed for texture analysis (EHD).

1.1.3 Shape Features

Because people can only distinguish between items based on their forms, shape characteristics give information for image retrieval and transmit situational data regarding a thing. One of the basic characteristics that helps us recognize items is shape. Using a region or a border as a starting point, shape extraction can be done. To extract shape characteristics, a variety of techniques like Fourier descriptor and moment invariants are used. Shape descriptors generally vary with scale and translation. To improve accuracy, they are typically combined with additional descriptors [4].

2 Literature Review

Researchers have proposed different methods to improve the system of content-based image retrieval.

Nazir et al. [5] developed a new CBIR technique is put forth for the mixing of colour and texture qualities. In order to extract colour information, the method uses a colour histogram (CH), while to extract texture features, it uses a discrete wavelet transform (DWT) and an edge histogram descriptor (EDH). Precision and recall techniques are used to produce results that are both competitive and effective. According to the experimental findings, the method put forth in this paper performs better than other CBIR systems already in use. absence of geographic data, lack of running-time data, and adoption of any machine learning technique. A VGGNet-based end-to-end CBIR was proposed by Zheng et al. [6].

They employed based similarity labels and data of universal gravitation in place of traditional labels to teach the network skills in feature extraction. Oxford Paris, Vacations, and Caltech 101 were the three datasets used for evaluation, and the suggested system achieved retrieval accuracy values of 0.9620, 0.9410, and 0.8850, respectively. Nevertheless, during the testing and training phases, the system requires major performance enhancements, as it takes a while to develop the curved spacetime database. To develop the database of general relativity, the pace of the testing and training phases must be enhanced over time. To develop the database of general relativity, the pace of the testing and training phases must be enhanced over time.

An information image retrieval technique that employs CNN to extract high-level characteristics from pictures was proposed by Sezavar et al. [7]. The computational cost is reduced by using the segmentation method, which is claimed to be useful in compressing. Both ALOI and Corel databases and MPEG-7 databases are used to evaluate the approach, and the results show that it has respectable retrieving speed and accuracy. Sparse representation guarantees quicker retrieval but with worse accuracy. The disadvantages are that fetching will be slower if dense encapsulation is not utilized, and the needed execution time will depend on how many photos are in the smaller sub.

Bani and Ershad [8], developed that the recommended image retrieval method extracts texture and colour information on a global and local scale in two spatial and frequency domains. Gray Level Co-occurrence matrices are generated in different directions after the image has been filtered using a Gaussian filter, and statistical features are then extracted. Compared to many current methods, the suggested method provides better precision. Some benefits of the suggested technique include rotation invariance, scale invariance, and low noise sensitivity. The drawbacks are do not bridge Semantic gap and high run time.

Latif and Khalil [9], examined the essential elements of several image retrieval and representation models, spanning from modern semantic deep-learning methods to low-level feature extraction. The features are extracted with the help of Color moment, Discrete Wavelet Transform (DWT), Gabor filter and Color and Edge Directivity Descriptor (CEDD). The effectiveness of a deep network on a huge unlabelled database in the context of unsupervised classification is thus one of the potential future study topics in this area. The drawbacks are high accuracy values and large feature vector dimension results in high computational expense.

Alsmadi et al. [10], In their proposed research, colour characteristics were extracted using YCbCr colour with discrete wavelet transform and Canny edge

histogram, shape features were extracted using RGB colour with neutrosophic clustering algorithm and Canny edge technique, and texture features were extracted using GLCM. The proposed CBIR system performed better than the existing methods and showed promising retrieval image results in terms of precision and recall rates. The proposed CBIR system has average recall and precision values of 90.15 and 18.03 respectively. The fact that convective cooling is so crucial as well as the longer calculation times are the downsides.

3 Proposed Methodology

The primary goal of this work is on content based image retrieval (CBIR) using a deep learning technique. The input images are acquired from the public database and features extracted using Histogram Oriented Gradient (HOG) descriptor. These features are applied to a Convolutional Neural Network (CNN) for accurate classification [11]. The methodology of proposed method is explained in detail with the help of the block diagram as shown in Fig. 1.

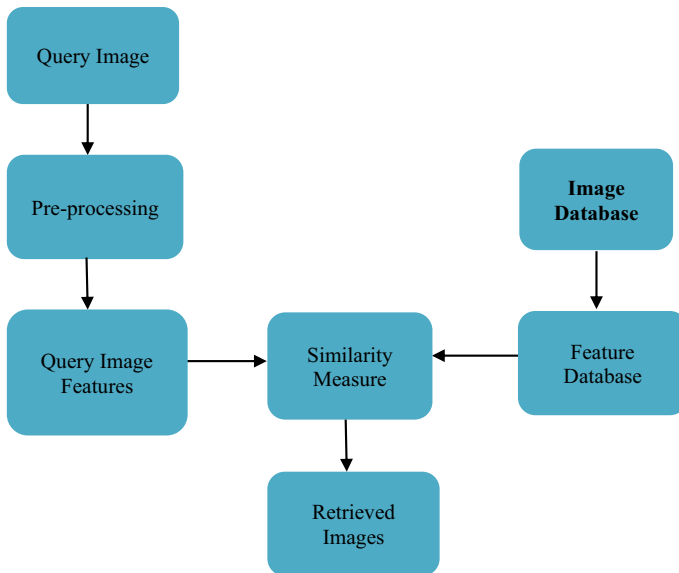


Fig. 1 Block diagram of proposed methodology

3.1 Query Image

Initially in the first step the user gives the query image and then that image is read by the system. The query image or input image taken in jpg format with different size from the database and then the query image is resized in the pre-processing step [12].

3.2 Classification

In this step the images which we take are resized and are converted into gray scale image. Here the query image is resized to 50×50 .

3.3 Feature Extraction

Features extraction is the process of storing the input data in the other format of the features. The variety of feature extraction approaches. Features namely colour, shape, texture and spatial information [13, 14]. In this step the features example colour is extracted from the query image as well as from the images in the image database. HOG descriptor features were used for extracting shape features. Hue, Saturation and Value (HSV) and Colour Histogram are used for extracting colour features. LBP image descriptor is used for extracting texture feature [15].

A feature descriptor called Histogram of Oriented Gradients (HOG), is frequently employed to extract features from image database [16]. It is commonly used for object detection in computer vision tasks. The HOG descriptor concentrates on the structure or shape of an image.

3.4 Image Database

An organized set of digital images aimed at the efficient management and the processing of queries on this image collection. In this project we took CBIR database consisting of 19 categories and each category contains one hundred images [17]. The CBIR dataset is of the following semantic categories: Bears, Buses, Horses, Flowers, Cats, Tigers, Tulips, Waterfalls, Dogs and Castles etc.

3.5 Similarity Measure

In this step the features which we extracted from the query image and the images in the image database are compared with the help of Convolutional Neural Network [18]. For deep learning algorithms, a CNN is a specific type of network architecture that is utilized for tasks like image recognition and processing similarity measures. It offers several attributes including adaptability, simple structure and less training parameters [19].

The layers of a convolutional neural network are numerous. To check and identify features that specifically reflect the input item, the complexity of the filters increases with each additional layer [20]. As a result, the output of each layer, also known as the convolved image, serves as the input for the next layer. In the last layer, the CNN recognizes the relevant images [21].

3.6 Retrieved Images

Finally, after finding or retrieving the relevant images from the previous step named similarity measure we get the retrieved images as output [22].

4 Simulation Results

In this project the proposed system for content-depends on image retrieval was evaluated using a number of query images and related images that were found in the database of images. Many query photos and similar images that were found in the database images were used to evaluate the content-based image retrieval technique system that was proposed.

In this project for experimental results initially an input image is taken from the category named Bears shown in Fig. 2. After comparing if the similarity between them is maximum then we get a dialogue box as shown in Fig. 4. Along with the dialogue box we get the relevant retrieved images as shown in Fig. 3.

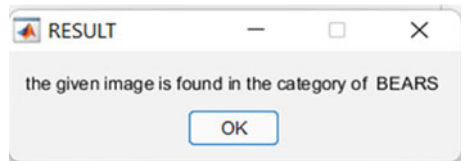
Fig. 2 Input image





Fig. 3 Retrieved images

Fig. 4 Dialogue box



In this project for experimental results initially an input image is taken from the category named Arabian Horses shown in Fig. 5. Features are obtained and analyzed from the input image and the database images. After comparing if the similarity between them is maximum then we get a dialogue box as shown in Fig. 7. Along with the dialogue box we get the relevant retrieved images as shown in Fig. 6.

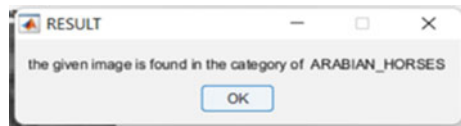
Fig. 5 Input image





Fig. 6 Retrieved images

Fig. 7 Dialogue box



5 Conclusion

In this project we discussed about a brand new CBIR technique and Image classification with the help of different image descriptors. Here we retrieved the relevant images from the image database with the help of different descriptors. From the query image and the images in the database features are extracted by using HOG and Gabor Wavelet.

After comparing the two color histogram features as well as comparing color and texture features, the project implemented a CBIR system using color and texture fused features. Similar images can be retrieved quickly and accurately by inputting a query image. On the basis of the input images, CNN and Gabor Wavelet are utilised to get images that are similar. The proposed CBIR system outperformed the already existing techniques and displayed encouraging retrieval image outcomes in terms of retrieval of images from various image collections.

References

1. Yellamma P, Sandeep PG, Revanth Sai R, Rohith Reddy S, Mahesh D (2023) Automatic vehicle alert and accident detection system based on cloud using IoT. In: *Embracing machines and humanity through cognitive computing and IoT*. Springer Nature Singapore, Singapore, pp 77–85
2. Guru Akhil T, Pranay Krishna Y, Gangireddy C, Kumar AK, Sowjanya KL (2022) Cyber-hacking breaches for demonstrating and forecasting. In: *Confidential computing: hardware based memory protection*. Springer Nature Singapore, Singapore, pp 169–175
3. Gaddam DKR, Ansari MD, Vuppala S, Gunjan VK, Sati MM (2022) A performance comparison of optimization algorithms on a generated dataset. In: *ICDSMLA 2020: Proceedings of the 2nd international conference on data science, machine learning and applications*. Springer Singapore, pp 1407–1415
4. Talukdar J, Singh TP, Barman B (2023) Time-series analysis. In: *Artificial intelligence in healthcare industry*. Springer Nature Singapore, Singapore, pp 109–126
5. Nazir A, Ashraf R, Hamdani T, Ali N (2018) Content based image retrieval system by using HSV color histogram, discrete wavelet transform and edge histogram descriptor. In: *International conferences on computing, mathematics and engineering technologies (iCoMET) 2(18):1–6*
6. Zheng Q, Singhal AM, Lall B (2019) Multi-channel local ternary pattern for content-based image retrieval. *Pattern Anal Appl* 22(8):1585–1596
7. Sezavar PD (2019) A new flower classification system using LBP and SURF features. *Int J Image Graph Signal Process* 11(5):13–20
8. Bani NT, Ershad SF (2019) Content-based image retrieval based on combination of texture and colour information extracted in spatial and frequency domains. *Electron Lib* 37(4):650–666
9. Latif A, Khalil T (2019) Content-based image retrieval and feature extraction: a comprehensive review. *Math Prob Eng* 12:23–29
10. Alsmadi MK (2020) Content-based image retrieval using color, shape and texture descriptors and features. *Arab J Sci Eng* 45(4):3317–3330
11. Gunjan VK, Kumar S, Ansari MD, Vijayalata Y (2022) Prediction of agriculture yields using machine learning algorithms. In: *Proceedings of the 2nd international conference on recent trends in machine learning, IoT, smart cities and applications: ICMISC 2021*. Springer Singapore, pp 17–26
12. Krishnaveni S, Harsha Priya M, Mallesh S, Narendar C (2022) A smart security systems using National Instruments myRIO. In: *Garcia Diaz V, Rincón Aponte GJ (eds) Confidential computing. advanced technologies and societal change*. Springer, Singapore. https://doi.org/10.1007/978-981-19-3045-4_20
13. Siddiquee KNEA, Islam MS, Singh N, Gunjan VK, Yong WH, Huda MN, Naik DB (2022) Development of algorithms for an IoT-based smart agriculture monitoring system. *Wirel Commun Mob Comput* 2022:1–16
14. Talukdar J, Singh TP, Barman B (2023) Methods of machine learning. In: *Artificial intelligence in healthcare industry. Advanced technologies and societal change*. Springer, Singapore. https://doi.org/10.1007/978-981-99-3157-6_3
15. Rashid E, Ansari MD, Gunjan VK, Ahmed M (2020) Improvement in extended object tracking with the vision-based algorithm. In: *Modern approaches in machine learning and cognitive science: a walkthrough: latest trends in AI*, pp 237–245
16. Anandita Iyer A, Umadevi KS (2023) Role of AI and its impact on the development of cyber security applications. In: *Artificial intelligence and cyber security in Industry 4.0*. Springer Nature Singapore, Singapore, pp 23–46
17. Lakshmana K, Shaik F, Gunjan VK, Singh N, Kumar G, Shafi RM (2022) Perimeter degree technique for the reduction of routing congestion during placement in physical design of VLSI circuits. *Complexity* 2022:1–11
18. Manasa K, Leo Joseph LMI (2023) IoT security vulnerabilities and defensive measures in Industry 4.0. In: *Sarveshwaran V, Chen JIZ, Pelusi D (eds) Artificial intelligence and cyber*

- security in Industry 4.0. Advanced technologies and societal change. Springer, Singapore. https://doi.org/10.1007/978-981-99-2115-7_4
19. Rudra Kumar M, Pathak R, Gunjan VK (2022) Machine learning-based project resource allocation fitment analysis system (ML-PRAFS). In: Computational intelligence in machine learning: select proceedings of ICCIML 2021. Springer Nature Singapore, Singapore, pp 1–14
 20. Merugu S, Kumar A, Ghinea G (2022) Geriatric mobility assistive system. In: Track and trace management system for Dementia and intellectual disabilities. Singapore: Springer Nature Singapore, pp 85–93
 21. Prasad PS, Sunitha Devi B, Janga Reddy M, Gunjan VK (2019) A survey of fingerprint recognition systems and their applications. In: ICCCE 2018: Proceedings of the international conference on communications and cyber physical engineering 2018. Springer Singapore, pp 513–520
 22. Merugu S, Kumar A, Ghinea G (2022) Real-time health monitoring using wearable devices. In: Track and trace management system for dementia and intellectual disabilities. Springer Nature Singapore, Singapore, pp 119–122