International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(3s) 2026 pp. 476-481
https://doi.org/10.70917/ijcisim-2026-2344
Article

Content Based Visual Information Retrieval using Deep Learning

Paresh Gagwani[1], Vinit Kotak[2]

[1] Department of Electronics Engineering, Shah & Anchor Kutchhi Engineering College, Mumbai, Maharashtra, India.
Email: pareshgagwani@gmail.com
[2] Department of Information Technology, Shah & Anchor Kutchhi Engineering College, Mumbai, Maharashtra, India.
Email: vinit.kotak@sakec.ac.in

Abstract: Content-based visual information retrieval (CBVIR) is a multidisciplinary area of research that focuses on using visual content such as images and videos for efficient and relevant information retrieval from large databases. Traditional text-based search techniques rely on metadata like keywords or tags, but CBVIR uses features extracted directly from the content itself—such as colors, textures, shapes, and patterns—allowing for more flexible and accurate searches. This paper reviews key advancements in CBVIR, including its techniques, challenges, and applications, and provides an overview of current research and trends in the field.

Keywords: NA

1. INTRODUCTION

Content-Based Visual Information Retrieval (CBVIR), also known as Content-Based Image Retrieval (CBIR), refers to the process of searching and retrieving digital images from large databases based on their visual content rather than relying solely on metadata such as keywords or descriptions. The exponential growth of digital content has made it increasingly difficult to manage and retrieve relevant data from large repositories. Traditional text-based retrieval systems, which rely on manually added keywords, are limited

by their inability to fully capture the rich, complex information present in visual content. In contrast, content-based visual information retrieval (CBVIR) systems aim to directly extract features from the image or video data, enabling more accurate and meaningful searches. This paper provides a comprehensive review of CBVIR techniques, highlighting the evolution of the field, various challenges, and their practical applications. This approach analyses the intrinsic features of images—such as colors, shapes, textures, and spatial arrangements—to facilitate efficient and accurate retrieval.

CBVIR systems operate by extracting low- level visual features from images and representing them in a feature space. These features are then indexed to facilitate efficient searching. A pivotal concept in CBVIR is the "visual word," which refers to distinctive image patches characterized by specific features like color, shape, or texture. By clustering similar visual words, systems can effectively organize and retrieve images based on visual similarity.

Despite advancements, CBVIR faces challenges such as:

- Semantic Gap: The disparity between low- level visual features and high-level semantic concepts.
- Scalability: Efficiently managing and searching through vast image databases.
- User Interaction: Developing intuitive interfaces that align with user expectations.

Future research in CBVIR is likely to focus on bridging the semantic gap through advanced machine learning techniques, improving scalability with optimized indexing methods, and enhancing user interaction by incorporating feedback mechanisms and personalized search options.

2. CBVIR USING MACHINE LEARNING

With the advent of machine learning, CBVIR systems began using more advanced feature extraction methods, including:

- SIFT (Scale Invariant Feature Transform) and SURF (Speeded-Up Robust Features): These algorithms detect and describe distinctive local features in images, which are invariant to scale and rotation.
- Bag of Visual Words (BoVW): This technique represents an image as a collection of visual words, which are obtained by quantizing feature descriptors into a dictionary.
- Deep learning: Convolutional neural networks (CNNs) and other deep learning models have significantly improved CBVIR by enabling automatic feature extraction directly from raw image data, reducing the need for hand-crafted features.

3. CORE TECHNIQUES IN CBVIR

CBVIR relies on several key techniques for processing visual content:

3.1 *Feature Extraction*

Feature extraction is the process of identifying key attributes of an image that can represent its content. The choice of features plays a crucial role in the accuracy and performance of CBVIR systems. Common feature types include:

- *Color-based features*: Histograms, color moments, and color correlograms.
- *Texture-based features*: Statistical methods, Gabor filters, and fractal-based features.
- *Shape-based features*: Edge histograms, Fourier descriptors, and region-based descriptors.

3.2 *Feature Matching and Indexing*

Once features are extracted, they need to be matched against a database of images to retrieve relevant results. Methods of matching include:

- *Euclidean distance*: A common method for comparing features by calculating the distance between feature vectors.
- *Similarity measures*: Techniques such as cosine similarity or correlation are used to assess the similarity between query and database images.
- *Indexing structures*: To efficiently store and retrieve large image databases, spatial indexing methods like KD-trees, R-trees, or hashing techniques are used.

3.3 *Retrieval Techniques*

Several retrieval techniques are used in CBVIR, including:

- *Query-by-Example (QBE):* The user provides an example image as a query, and the system retrieves similar images.
- *Query-by-Sketch (QBS):* The user provides a simple sketch of the desired image, and the system retrieves images based on the sketch.
- *Semantic-based retrieval*: By incorporating machine learning and deep learning, systems can recognize high-level concepts (e.g., identifying objects, people, or scenes) and perform more intuitive searches.

3.4 *Relevance Feedback*

Relevance feedback is a mechanism that allows users to refine their queries based on the results they receive. The user indicates which results are relevant, and the system uses this feedback to adjust the ranking of images in subsequent retrievals. This is often done using techniques like:

- Supervised learning: Refining the model based on user feedback.
- Active learning: The system selects the most informative images to request feedback on.

4. RELATED WORK

Divya Shrivastava et al [1], presented a rigorous survey on Local and Global Features Selection, Extraction, Representation, and Evaluation parameters for Content Based Image Retrieval. The authors also discussed the Taxonomy of Deep Learning-Based Image Retrieval Methods.

Sarath Chandra Yenigalla et al [2], described their research work which is the Implementation of Content-Based Image Retrieval Using Artificial Neural Networks. In their research work the authors examine the CBIR system utilizing three machine learning methods, namely SVM (Support Vector Machine), KNN (K Nearest Neighbours), and CNN (Convolution Neural Networks), using Corel 1K, 5K, and 10K databases, by splitting the data into 80% train data and 20% test data. Moreover, compare each algorithm's accuracy and efficiency when a specific task of image retrieval is given to it. The final outcome of this project will provide us with a clear vision of how effective deep learning, KNN and CNN algorithms are to finish the task of image retrieval.

Manoharan Subramanian et al [3], proposed a new approach for Content-Based Image Retrieval (CBIR) has been addressed by extracting colour, gray, advanced texture, and shape features for input query images. Contour-based shape feature extraction methods and image moment extraction techniques are used to extract the shape features and shape invariant features. The informative features are selected from extracted features and combined colour, gray, texture, and shape feature s by using PSO. The target image has been retrieved for the

given query image by training the random forest classifier. The proposed colour, gray, advanced texture, shape feature, and random forest classifier with optimized PSO (CGATSFRFOPSO) provide efficient retrieval of images in a large-scale database. The main objective of this research work is to improve the efficiency and effectiveness of the CBIR system by extracting the features like colour, gray, texture, and shape from database images and query images. These extracted features are processed in various levels like removing redundancy by optimal feature selection and fusion by optimal weighted linear combination. The Particle Swarm Optimization algorithm is used for selecting the informative features from gray and colour and texture features. The matching accuracy and the speed of image retrieval are improved by an ensemble of machine-learning algorithms for the similarity search.

Rabia Younas et al [4], presented an implementation of a CBIR framework that not only tries to efficiently capture the user intent based on the feedback but also provides query suggestions that can help its users to pose better queries to retrieve desired results efficiently. The proposed technique is straightforward to implement and scope efficiently to huge datasets. Extensive experiments on diverse real datasets with image similarity measures have revealed the dominance of the proposed method over original algorithms.

Xiuqu Li et al [5] presented an effective content-based visual image retrieval system. his system consists of two main components: visual content extraction and indexing, and query engine. Each image in the image database is represented by its visual features: color and spatial information. The system uses a color label histogram with only thirteen bins to extract the color information from an image in the image database. A unique unsupervised segmentation algorithm combined with the

wavelet technique generates the spatial feature of an image automatically. The resulting feature vectors are relatively low in dimensions compared to those in other systems. The query engine employs a color filter and a spatial filter to dramatically reduce the search range. As a result, queue processing is speeded up.

Luca Piras et al [6], presented a journey through the main information fusion ingredients that a recipe for the design of a CBIR system should include to meet the demanding needs of users. Their contribution in the research work include Feature weighting for early fusion, Representation by multi-feature spaces for late fusion, Fusing different relevance feedback approaches and Multimodal retrieval.

5. PROPOSED SYSTEM

We propose to derive feature vectors of an image with the help of some pre-trained deep- learning models.. To evaluate the CBVIR model, the combination of various quantifiable qualities for its three modules VFE, VFL and VMM compose a complex evaluation process. Metrics from the respective domains of Deep Neural Networks, Data Mining and Information Retrieval are used for individual evaluation of the CNN model, clustering model, and IR model.

An additional factor of difficulty for optimizing the performance of the pipeline is the choice of data. There are two image subsets for training and evaluating a CNN classifier.

A part of them can be indexed by the CBVIR engine as the image collection, while some images will remain unknown to the system in order to be used for the test queries. The retrieved set of images will provide an indication for the performance of the overall system.

Fig. 1. Sample images from Dataset

(a) ImageDB2000 and (b) DBCaltech.

Block diagram of the Content based Visual Information Retrieval is shown in figure 2. In this system the Convolutional Neural Network (CNN) which is a deep learning model is used.

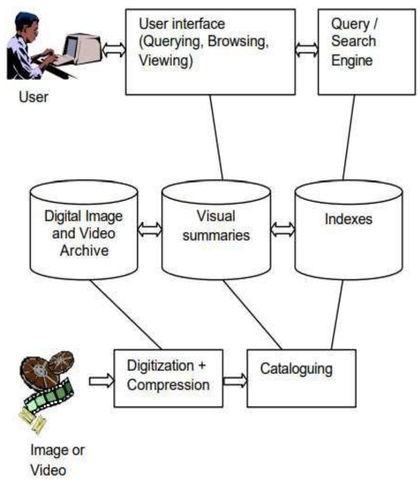

Figure 2. Block diagram of a CBVIR system.

The proposed Content based Visual Information Retrieval System is based on Convolutional Neural network which has multiple layers. The feature representation process is depicted as below in figure 3.

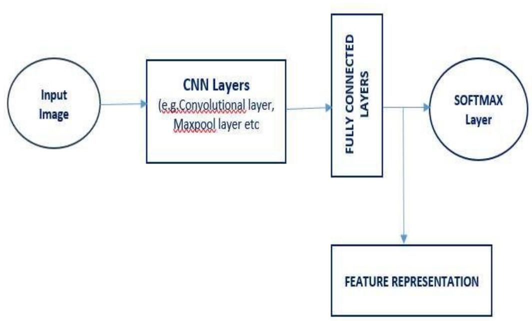

Fig. 3: Block diagram of CBVIR image feature representation with the help of pre-trained deep-learning models.

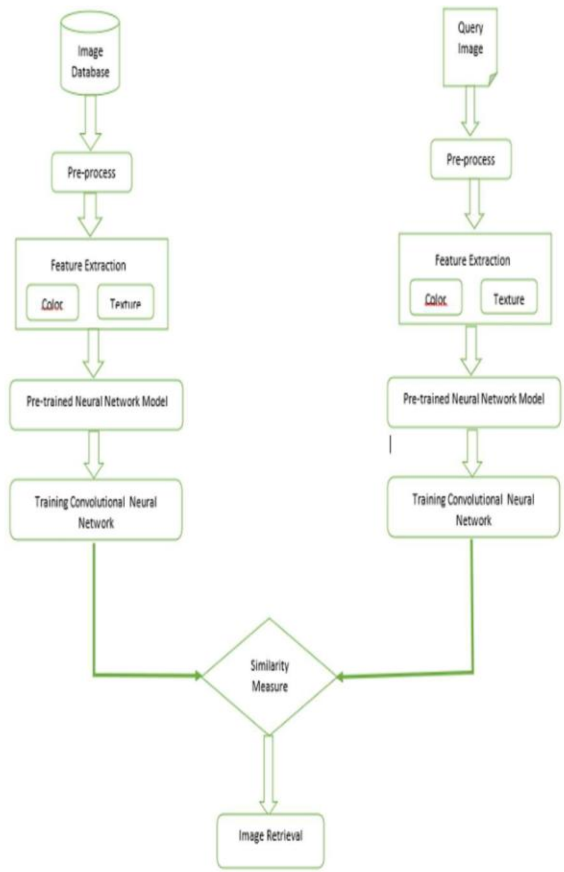

Fig 3: Flowchart of the proposed CBVIR system

6. RECENT TRENDS AND FUTURE DIRECTIONS

Recent research in CBVIR focuses on the following areas:

- Deep learning and CNNs: Further advancements in deep learning models continue to improve the accuracy of feature extraction and retrieval.
- Multimodal retrieval: Combining visual data with other modalities, such as text and audio, to improve search results.
- Transfer learning: Leveraging pre-trained models on large datasets to improve CBVIR performance on smaller, domain-specific datasets.
- Explainability and interpretability:

Developing systems that can explain why

a certain image was retrieved or how it relates to the query.

7. CONCLUSION

In this paper, we provided an up-to-date review of content-based visual information retrieval systems. Content-based visual information retrieval has made significant strides in recent years, offering more accurate and efficient ways to search and retrieve visual data. While challenges such as the semantic gap and scalability remain, advancements in machine learning, especially deep learning, hold promise for further improving the performance of CBVIR systems. The future of CBVIR lies in integrating it with other modalities, enhancing user interaction, and addressing computational and interpretability challenges.

References

1. Divya Shrivastava, Shashank Sheshar Singh,

2. B. Rajitha, Madhushi Verma, Manjit Kaur, Heung-No, “Content-Based Image Retrieval: A Survey on Local and Global Features Selection, Extraction, Representation, and Evaluation Parameters”, IEEE Access, Vol. 11, August 2023
3. Sarath Chandra Yenigalla *, Karumuri Srinivasa Rao and Phalguni Singh Ngangbam, “Implementation of Content-Based Image Retrieval Using Artificial Neural Networks”,
4. MDPI, Engineering Proceedings, March 2023
5. Manoharan Subramanian,Velmurugan, Lingamuthu ,Chandran Venkatesan,
6. and Sasikumar Perumal, “Content-Based Image Retrieval Using Colour, Gray, AdvancedTexture, Shape Features, and Random Forest Classifier with optimized Particle Swarm Optimization”, Hindawi International Journal of Biomedical Imaging, 2022
7. Rabia Younas, Hafiz Burhan Ul Haq, Muhammad Daniyal Baig, “A Framework for Extensive Content-Based Image Retrieval System Incorporating Relevance Feedback and Query Suggestion”, Spectrum of Operational Research, Volume 1, Issue 1 (2024
8. Subhadip Maji, Smarajit Bose. "CBIR Using Features Derived by Deep Learning" , ACM/IMSTransactions on Data Science, 2021
9. Khawaja Ahmed, Shahida, and Muhammad Iqbal. 2018. Content-based image retrieval using image features information fusion. Info. Fusion 51 (Nov. 2018)
10. K. T. Ahmed, S. A. H. Naqvi, A. Rehman, and T. Saba. 2019. Convolution, approximation and spatial nformation- b a s e d object and color signatures for content based image retrieval. In Proceedings of the International Conference on Computer and Information Sciences (ICCIS’19). 1–6.
11. Aksoy and R. M. Haralick. 2000. Probabilistic vs. geometric similarity measures for image retrieval. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR’00), Vol. 2. 357–362.
12. Mutasem K. Alsmadi. 2020. Content-based image retrieval using color, shape and texture descriptors and features. Arab. J. Sci. Eng. 45, 4 (2020), 3317–3330
13. Xiuqu Li, Shu-Ching Chen, Mei Ling Shyu, B. Furht, “An effective content-based visual image retrieval system”, IEEE Explore 2002
14. Luca Piras, Giorgio Giacinto, “Informatio fusion in content based image retrieval: A comprehensive overview”, Information Fusion, ScienceDirect, Elsevier, Vol .37, 2017