

Recent Advances in Content-based Image Retrieval: Techniques and Applications

Mahesh Prajapati^{1*}, Vijay Bhandari², Ritu Shrivastava²

Abstract

Content-based image retrieval (CBIR) plays a vital role in computer vision, driven by the increasing need for fast and accurate image retrieval across fields like healthcare, e-commerce, and digital libraries. This study offers a detailed review of CBIR methodologies, charting their progression from traditional feature extraction techniques, such as Local Binary Patterns (LBP), to contemporary deep learning-driven methods. The transformative impact of convolution neural networks (CNNs) is highlighted, emphasizing their ability to capture intricate, hierarchical features that address the semantic gap and enhance retrieval precision. The study evaluates various algorithms, examining their strengths and weaknesses while tackling key challenges like computational efficiency, scalability for large datasets, and the enduring semantic gap in interpreting image content. It explores emerging trends, including hybrid feature models, attention mechanisms, multi-modal learning, and transfer learning, showcasing their potential to overcome existing hurdles. Privacy-preserving methods like federated learning and domain-specific solutions, particularly in medical imaging, are also discussed for their practical relevance. The real-world use of CBIR systems spans industries, from healthcare for diagnostic support to e-commerce for personalized recommendations and cultural heritage preservation, illustrating the method's versatility. This review seeks to give researchers a comprehensive understanding of the CBIR field, highlighting critical research gaps and promising future directions. These include lightweight architectures for real-time applications, effective feature fusion techniques, and adaptive learning strategies tailored to specific domains.

Keywords: Content-based image retrieval (CBIR), feature extraction, deep learning, convolution neural networks (CNNs), semantic gap, application of CBIR

INTRODUCTION

Over the past few decades, Content-based Image Retrieval (CBIR) has transitioned from an academic concept to a foundational component of modern computer vision applications. Its ability to retrieve images based on visual content rather than metadata has unlocked new possibilities across diverse fields

such as medical imaging, e-commerce, and digital libraries. This study provides an in-depth review of recent advancements in CBIR, with a particular emphasis on the evolution of feature extraction techniques and the adoption of deep learning models.

CBIR systems have significantly advanced, evolving from simple feature-based approaches to highly complex deep learning frameworks. Early methods used handcrafted features such as SIFT or HOG, which, despite their effectiveness, often struggled to bridge the semantic gap between low-level features and high-level concepts. The introduction of deep learning, particularly CNNs, has revolutionized CBIR by enabling end-to-end

*Author for Correspondence

Mahesh Prajapati
E-mail: tguy68086@gmail.com

¹Student, Department of Computer Science and Engineering, Sagar Institute of Research and Technology, Bhopal, Madhya Pradesh, India

²Professor, Department of Computer Science and Engineering, Sagar Institute of Research and Technology, Bhopal, Madhya Pradesh, India

Received Date: December 04, 2024

Accepted Date: December 19, 2024

Published Date: December 24, 2024

Citation: Mahesh Prajapati, Vijay Bhandari, Ritu Shrivastava. Recent Advances in Content-based Image Retrieval: Techniques and Applications. Journal of Computer Technology & Applications. 2025; 16(1): 67–71p.

feature learning and extraction. This study reviews these developments, emphasizing hybrid models, scalable indexing techniques, and practical real-world applications [1].

CLASSIFICATION METHODS

CBIR employs various classification approaches, each offering unique benefits and limitations.

Traditional Feature-based Approaches

Traditional methods utilize handcrafted features such as:

- *SIFT (Scale-Invariant Feature Transform)*: Effective for local feature matching.
- *HOG (Histogram of Oriented Gradients)*: Useful for texture and shape analysis.

CNN-based Methods

CNNs extract hierarchical features, enabling improved semantic understanding. Pre-trained models like ResNet, VGG16, and Efficient Net have achieved high retrieval accuracy by leveraging transfer learning [2].

Attention Mechanisms and Transformers

Recent developments integrate attention layers to enhance feature extraction by focusing on important regions within an image. Transformers have also emerged as powerful tools for image retrieval tasks [3].

- *Advantages*: Improved context understanding.
- *Disadvantages*: High computational requirements.

Hybrid Techniques

Combining traditional and deep learning features enhances retrieval accuracy while maintaining computational efficiency.

- *Example*: Hybrid models utilizing Local Binary Patterns (LBP) and CNN embeddings have shown significant promise.

Relevance Feedback Systems

Relevance feedback iteratively improves retrieval performance by incorporating user preferences. Modern systems integrate neural networks with relevance feedback mechanisms for better adaptability [4].

LITERATURE REVIEW

Here, we have studied the following research papers to understand the recent work on stock price prediction.

Attention-based Convolutional Neural Networks for Chinese Relation Extraction by Zhang *et al.* [3]

This paper introduces an attention-based CNN architecture designed to dynamically concentrate on key regions within images during feature extraction, enhancing retrieval accuracy in datasets with complex backgrounds. The method leverages attention mechanisms embedded within CNNs to facilitate region-specific feature learning. However, a significant drawback of this approach lies in its high computational demands, which hinder real-time performance and scalability for large datasets [3].

Scalable Deep Hashing for Large-scale Social Image Retrieval by Cui *et al.* [2]

This paper presents a hashing-based indexing approach tailored for deep feature representations, facilitating efficient retrieval in large-scale image databases. The method integrates deep embeddings with binary hash codes to enable rapid similarity searches. However, a key challenge is that hashing introduces slight precision losses in retrieval outcomes, particularly for images that are visually similar [2].

An Effective Hybrid Framework for Content-based Image Retrieval (CBIR)

by Khan *et al.* [5]

This paper introduces a hybrid CBIR framework that combines handcrafted descriptors like Local Binary Patterns (LBP) with deep learning embeddings to improve accuracy. The methodology integrates LBP with CNN-based features, harnessing the strengths of both techniques. However, the hybrid approach is computationally demanding due to the parallel processes involved in feature extraction and fusion [5].

Exploring Multi-modal Contextual Knowledge for Open-vocabulary Object Detection

by Xu *et al.* [1]

This paper tackles the semantic gap issue by combining visual and textual metadata through multi-modal deep learning, allowing retrieval systems to simultaneously account for contextual and semantic information. The proposed approach utilizes multi-modal encoders to integrate visual and textual features. However, its effectiveness depends heavily on the availability of accurate metadata annotations, which can often be incomplete or noisy in real-world datasets [1].

Content-based Medical Image Retrieval System for Lung Diseases Using Deep CNNs

by Agrawal *et al.* [6]

The authors developed a CBIR system for medical images, leveraging pre-trained CNNs for feature extraction and similarity matching. This approach enhanced diagnostic support by retrieving comparable medical cases. The method involved fine-tuning a ResNet model specifically for domain-focused image retrieval in healthcare. However, it encountered challenges with inter-class variations, particularly in datasets containing images that were visually similar but diagnostically distinct [6].

Recent Developments of Content-based Image Retrieval (CBIR)

by Li *et al.* [4]

This paper combines visual-based CBIR with recommendation systems to enhance e-commerce platforms, facilitating personalized image-based searches. The approach integrates product metadata with visual features using a neural collaborative filtering technique. However, it faces scalability challenges in dynamic environments where user preferences and product catalogs evolve rapidly [4].

Medical Image Analysis Using Wavelet Transform and Deep Belief Networks

by Khatami *et al.* [7]

This paper introduces a hybrid approach that combines wavelet transforms with CNNs to enhance retrieval efficiency in medical imaging. It utilizes wavelet decompositions for texture analysis, integrating them with CNN-based embeddings for a more robust feature representation. However, the system's generalizability is limited to medical datasets due to the specificity of the chosen feature set [7].

Few-shot Image Recognition with Knowledge Transfer

by Peng *et al.* [8]

This study demonstrated that transfer learning with pre-trained CNN models can significantly shorten training time while preserving high retrieval accuracy. The approach involved fine-tuning ImageNet-trained models for domain-specific CBIR tasks. However, the method is constrained by the quality of the pre-trained models and their applicability to specific niche domains [8].

Federated Learning and Differential Privacy for Medical Image Analysis

by Adnan *et al.* [9]

The authors introduced federated learning for CBIR to ensure data privacy while enabling collaborative model training across distributed datasets. A federated learning framework was developed to train CBIR systems without sharing raw images. However, the approach faces challenges such as computational overhead from synchronization and reduced accuracy compared to centralized training models [9].

Transformers in Medical Imaging: A Survey by Shamshad *et al.* [10]

This study examined the application of transformers in CBIR, emphasizing their capacity to model long-range dependencies in images. By integrating self-attention mechanisms, the method effectively captured global relationships within images. However, high computational complexity and memory demands constrained the use of transformers to smaller datasets [10].

PROPOSED METHOD

The proposed Content-based Image Retrieval (CBIR) system blends advanced deep learning techniques with an intuitive web interface for seamless image searching. It leverages a pre-trained convolution neural network to extract detailed, high-dimensional features from both the query and dataset images, encapsulating both semantic and visual elements. These features are organized in a structured database, allowing the system to compute cosine similarity between the query image's features and those in the dataset, efficiently ranking results by relevance. Here we will be using a full stack technology in which we will be creating a front end using technology such HTML, CSS, JavaScript to show the main interface and in the backend we will be using Django framework of python; the training data to be collected from the open source libraries and the training data to be stored in the database such as mysql or sqlite for the faster retrieval or feature storing. The Query image will be analyzed and features of both query and training data to be matched to get accurate data. The system workflow involves three main stages:

- *Dataset preparation*, where images are preprocessed and features are extracted.
- *Feature matching*, where a query image is compared against the dataset using similarity metrics.
- *Result visualization*, where the top N similar images are presented to the user. This approach addresses scalability and efficiency while minimizing the semantic gap by leveraging deep feature representations. Additionally, the web interface allows users to upload images seamlessly and retrieve results in real time as shown in Figure 1.

This method is scalable, adaptable to various domains, and capable of handling large datasets effectively, making it suitable for applications in medical imaging, e-commerce, and digital libraries.

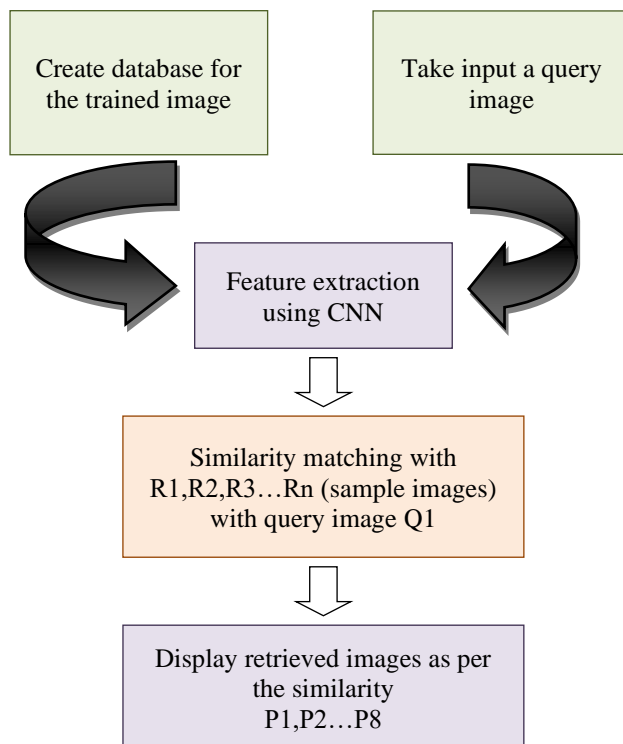


Figure 1. Proposed method flow diagram.

CONCLUSION

Based on the reviewed literature, there have been notable advancements and diverse methodologies in Content-based Image Retrieval (CBIR) systems. Initially reliant on handcrafted features such as Local Binary Patterns (LBP) and traditional convolutional neural networks (CNNs), contemporary techniques have shifted toward deep learning-based models and hybrid architecture. These approaches enhance retrieval accuracy and minimize the semantic gap. Features like attention mechanisms, multi-modal learning, and the integration of recommendation systems have demonstrated effectiveness in improving performance across fields such as medical imaging and e-commerce.

However, challenges persist. High computational demands for models like attention-based CNNs, accuracy trade-offs in hashing methods, and dependence on precise metadata for multi-modal learning remain significant barriers. Privacy-preserving methods such as federated learning address data security concerns but bring synchronization complexities and reduced accuracy compared to centralized systems.

To overcome these challenges, the proposed research method integrates deep learning-based feature extraction with scalable and efficient similarity matching in a robust database system. It employs a pretrained CNN for feature extraction and uses metrics like cosine similarity for precise, real-time image retrieval. The system includes a user-friendly web-based interface developed with Django or Flask, enabling smooth image uploads and retrievals. This approach balances efficiency, accuracy, and scalability, targeting practical applications in areas like medical imaging, e-commerce, and digital libraries while addressing gaps in existing solutions.

Acknowledgment

We wish to express my gratitude to my Institute for their unwavering guidance and support. We also appreciate SIRT providing research resources. Special thanks to those who shared insights, participated in surveys, and contributed to this work. We are thankful for the encouragement and understanding of my friends and family.

REFERENCES

1. Xu Y, Zhang M, Yang X, Xu C. Exploring multi-modal contextual knowledge for open-vocabulary object detection. *IEEE Trans Image Process.* 2024 Oct 29; 33: 6253–6267.
2. Cui H, Zhu L, Li J, Yang Y, Nie L. Scalable deep hashing for large-scale social image retrieval. *IEEE Trans Image Process.* 2019 Sep 16; 29: 1271–84.
3. Wu W, Chen Y, Xu J, Zhang Y. Attention-based convolutional neural networks for Chinese relation extraction. In *Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data: 17th China National Conference, CCL 2018, and 6th International Symposium, NLP-NABD 2018, Changsha, China, October 19–21, 2018, Proceedings 17*; Springer International Publishing. 2018; 147–158.
4. Li X, Yang J, Ma J. Recent developments of content-based image retrieval (CBIR). *Neurocomputing.* 2021 Sep 10; 452: 675–89.
5. Khan UA, Javed A, Ashraf R. An effective hybrid framework for content based image retrieval (CBIR). *Multimed Tools Appl.* 2021 Jul; 80(17): 26911–37.
6. Agrawal S, Chowdhary A, Agarwala S, Mayya V, Kamath SS. Content-based medical image retrieval system for lung diseases using deep CNNs. *Int J Inf Technol.* 2022 Dec; 14(7): 3619–27.
7. Khatami A, Khosravi A, Nguyen T, Lim CP, Nahavandi S. Medical image analysis using wavelet transform and deep belief networks. *Expert Syst Appl.* 2017 Nov 15; 86: 190–8.
8. Peng Z, Li Z, Zhang J, Li Y, Qi GJ, Tang J. Few-shot image recognition with knowledge transfer. In *Proceedings of the IEEE/CVF international conference on computer vision.* 2019; 441–449.
9. Adnan M, Kalra S, Cresswell JC, Taylor GW, Tizhoosh HR. Federated learning and differential privacy for medical image analysis. *Sci Rep.* 2022 Feb 4; 12(1): 1953.
10. Shamshad F, Khan S, Zamir SW, Khan MH, Hayat M, Khan FS, Fu H. Transformers in medical imaging: A survey. *Med Image Anal.* 2023 Aug 1; 88: 102802.