Content Based Image Retrieval

ECE501- Digital Image Processing
Mid-Sem Report - Group 6

Dhriti Gandhi (AU2340030), Aneri Kabrawala (AU2340041), Renee Vora (AU2340059), Pushti Sonak (AU2340082)
School of Engineering and Applied Science (SEAS)
Ahmedabad University, Ahmedabad, India

***Abstract*—The growing necessity for reliable image retrieval for practical applications has increased due to the abundance of images available across domains. Content-Based Image Retrieval (CBIR) is one of the methods for this purpose that retrieves images within a pool of datasets on the basis of the visual contents (features). This work presents a classical CBIR system for retrieving similar images using colour, edge, texture and fusion feature techniques. The effectiveness and accuracy of retrieval are used to evaluate the system's performance, which highlights the features' capacity to provide relevant search results.**

***Index Terms*—CBIR, feature fusion, colour histograms, Euclidean distance**

I. Introduction

In the current digital age, the internet is full of images of every domain, and so are our personal devices. It is quite challenging to retrieve and organise these images according to the practical applications. One of the traditional methods of image retrieval was through textual descriptions and assigned tags. However, such methods are ambiguous and give results that are subjective, incomplete and inaccurate. Content-Based Image Retrieval (CBIR), is a technique that solves this problem by using actual features of the image, like colours, textures, edges to search for images in a dataset. The extracted measurable features are compared within the set of images to find matches to retrieve similar images. CBIR has many applications, like effective image search, security and surveillance, medical imaging, and E-commerce.

This project tries to implement a CBIR system using classical image processing approaches. We have conducted a literature review to explore the existing feature extraction methods, their advantages and disadvantages. CBIR systems process a given dataset of images to extract meaningful features and retrieve similar images on the basis of a query based on these features. Here, the features used for retrieval are: colour, edge, texture and a fusion of colour and edge features.

II. Methodology

The flowchart given below depicts the flow of the program and algorithm we have been using till now. We start with a custom dataset which is currently a mixture of all types of images to check the program's effectiveness. Preprocessing is done on the dataset It includes additional padding on the images to make the images of the same size; resizing to a particular dimension; convert to HSV Color Space and also show Histogram Equalization. Additional padding and resizing makes sure that the images do not stretch out or seem too compressed. Next we choose a Query Image on which we perform Feature Extraction methods (Color ,Texture ,Edge). We have just done individual methods and Color-Edge fusion till now but aim to try out more feature fusion methods with different retrieval methods.

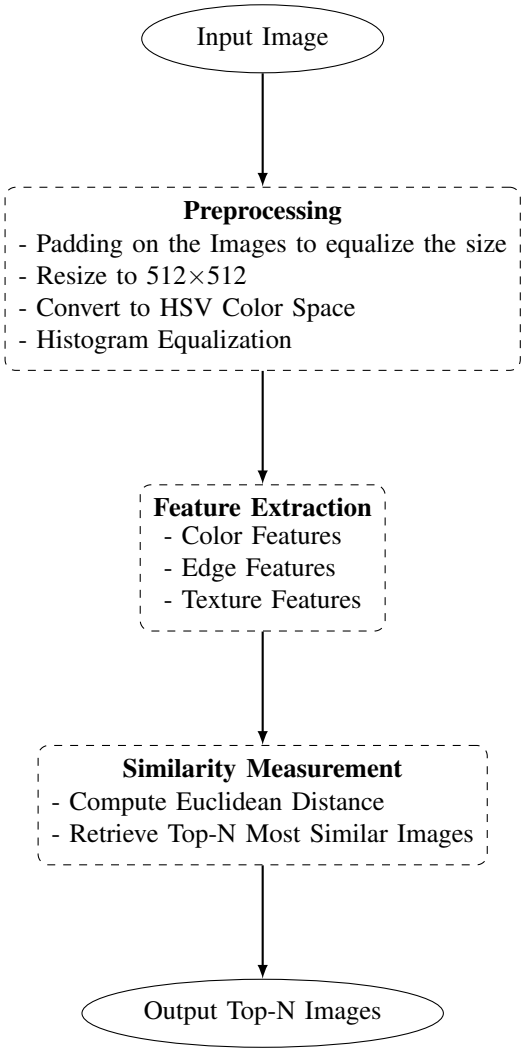


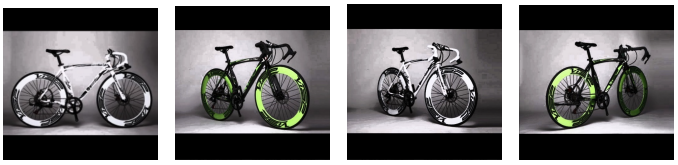
Fig. 1: Vertical flowchart of image retrieval process.

Finally, the system computes the euclidean distance between

the feature vectors of the query image andthe dataset. Based on these similarity scores, it will retrieve and display the Top-N most simial images as output. This process forms the foundation of our Content Based Image Retrieval (CBIR) framework, which combines classical image processing and feature-based comparison for visual similarity detection.

III. Results

Color-based Extraction:



(a) Query (b) Rank 1 (c) Rank 2 (d) Rank 3

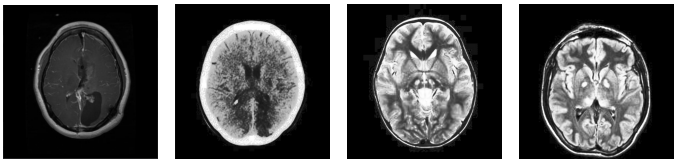
Fig. 2: Color-based Extraction Results

Edge-based Extraction:

(a) Query (b) Rank 1 (c) Rank 2 (d) Rank 3

Fig. 3: Edge-based Extraction Results

Color- & -Edge based Extraction:



(a) Query (b) Rank 1 (c) Rank 2 (d) Rank 3

Fig. 4: Color- & Edge-based Extraction Results

IV. Discussions

The Content-Based Image Retrieval system demonstrates how the color, edge, and texture features extraction help to retrieve similar images. In color feature extraction, using the HSV histogram, the results show the dominant color retrieved images, with the first result showing yellow and green, similar to the query image. For edge feature extraction, the technique used is the Sobel edge detection, the results of the query image have similar outlines or shapes irrespective of the color. The fusion of the above two features, i.e., fusion of color histograms and Sobel edge detection, improves the retrieval performance compared to individual features, as seen for the MRI images. The results from the texture feature extraction, which uses the Local binary pattern, give less accurate results that need to be improved by using other texture feature extraction techniques like GLCM, mainly for MRI images.

V. Conclusion

We have successfully completed our first phase of experimenting that is individual feature extraction along with feature fusion with Color-Edge. In next phase, we will focus on implementing and testing all three feature-fusion along with weighted fusion. Trial of different retrieval algorithms is also on the next phase of tasks. Finally , we will centralize our dataset to a specific domain, incorporate domain-specific image properties, and conduct performance analysis against existing models for comprehensive evaluation.

References

[1] C. Knoblock, "An example of Content-Based Image Retrieval," *ResearchGate*, [Online image]. Available: https://www.researchgate.net/profile/Craig-Knoblock/publication/221607221/figure/fig3/AS:305582780239874@1449868020143/An-example-of-Content-Based-Image-Retrieval.png. [Accessed: Oct. 12, 2025].

[2] H. Farsi and S. Mohamadzadeh, "Colour and texture feature-based image retrieval by using Hadamard matrix in discrete wavelet transform," *IET Image Processing*, vol. 7, no. 3, pp. 212–218, Apr. 2013. doi: https://doi.org/10.1049/iet-ipr.2012.0203.

[3] K. Haridas and A. Selvadoss Thanamani, "Efficient Content Based Image Retrieval System in Visual Words, Color and Edge Directive Descriptors and Fuzzy Color and Texture Histogram," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 3297, 2007. [Online]. Available: https://www.rroij.com/open-access/efficient-content-based-image-retrievalsystem-in-visual-words-color-and-edgedirective-descriptors-and-fuzzy-color-andtexture-histogram.pdfrroij.com/fuzzy-color-texture-histogram. [Accessed: Oct. 12, 2025].

[4] M. Singha and K. Hemachandran, "Content Based Image Retrieval using Color and Texture," *Signal & Image Processing: An International Journal*, vol. 3, no. 1, pp. 39–57, Feb. 2012. doi: https://doi.org/10.5121/SIPIJ.2012.3104. Available: https://aircconline.com/sipij/V3N1/3112sipij04.pdf.

[5] R. Kumar and N. Murthy M. S., "Enhanced Content Based Image Retrieval Using Integrated Color and Texture Features," *International Journal on Science and Technology*, vol. 16, no. 1, Jan. 2025. doi: https://doi.org/10.71097/ijsat.v16.i1.1418.

[6] A. Bajaj, "CBIR_dataset [Data set]," *Kaggle*, 2020. Available: https://www.kaggle.com/datasets/theaayushbajaj/cbir-dataset.

[7] M. Ache, "GPR1200 Dataset," 2021. [Online]. Available: https://www.kaggle.com/datasets/mathurinache/gpr1200-dataset/data. [Accessed: Oct. 12, 2025].

[8] M. Nickparvar, "Brain Tumor MRI Dataset [Data set]," *Kaggle*, 2021. doi: https://doi.org/10.34740/KAGGLE/DSV/2645886.

[9] Visual Geometry Group, "Oxford 17 Flowers Dataset," 2008. [Online]. Available: https://www.robots.ox.ac.uk/~vgg/data/flowers/17/index.html. [Accessed: Oct. 12, 2025].