Ahmedabad University

ECE501- Digital Image Processing

Weekly Report 5

Section 1

Submitted to faculty: Prof. Mehul Raval

Topic: Content-Based Image Retrieval (CBIR)

Week Duration: 25th Oct - 1st Nov

Group no.: 6

Enrollment No.	Name	Name of the Program
AU2340030	Dhriti Gandhi	Btech: Computer Science And Engineering
AU2340041	Aneri Kabrawala	Btech: Computer Science And Engineering
AU2340059	Renee Vora	Btech: Computer Science And Engineering
AU2340082	Pushti Sonak	Btech: Computer Science And Engineering

2025-2026 (Monsoon Semester)

Objective

To design a Content-Based Image Retrieval (CBIR) system capable of retrieving visually similar images from a custom dataset using classical Digital Image Processing techniques such as colour histograms, texture analysis and edge detection, test and verify the results on images from the collected dataset, and retrieve meaningful data.

Input: a query image and a dataset
Output: a set of images from the dataset with a rank which are visually similar to the input image.

Work Done This Week

- Researched on digital libraries for drawings dataset to explore the kind of features required for the same in CBIR.
- We thoroughly discussed on what exact type of dataset to choose within the drawings domain (digital paintings, sketches, indian paintings, professional art, etc.).
- We decided sub-domains in each domain by reviewing what methods would be needed for each.
- We revised our dataset for all three domains. For medical scans - we took MRI scans, mainly of brain. For natural sceneries- we mainly took flora and fauna centered images. And finally, for drawings, we took both sketches and Indian paintings (from kaggle).
- We researched and decided key metrics for evaluation purposes.

Next Week Plan

- Plan to apply domain specific features for at least 2 domains.
- Get the performance metrics for each domain and also compare with the baseline research.
- Find more similar papers and compare our work with them.

Challenges Faced

- We found upon our research, that the digital libraries dataset required extensive ML/DL for accurate outputs. Also the data was too random and was of vast range with varied sub-domains. This did not aligned to our project's requirements.
- Also, difficult to find open and easily accessible dataset in digital libraries. The dataset was not available as whole and had t be downloaded individually or run an API code in colab for the same.

- In general, the run time while testing the code was too long, which tested our patience and slowed down our progress.