Metadata-Augmented Image Indexing and Search Using FAISS

Similar Image Search using FAISS and ResNet Feature Extraction

the project implements find similar image in the dataset for the query image and if the image can't be matched find the similar based on the rating to the query image

This project implements a FAISS-based HNSW (Hierarchical Navigable Small World) indexing system for fast and efficient similar image search using ResNet-based feature extraction.

It also incorporates PCA for dimensionality reduction and metadata analysis using image EXIF data.

Search

The Query image existing in the data set and find top 2 similar images

Query image

search image 1

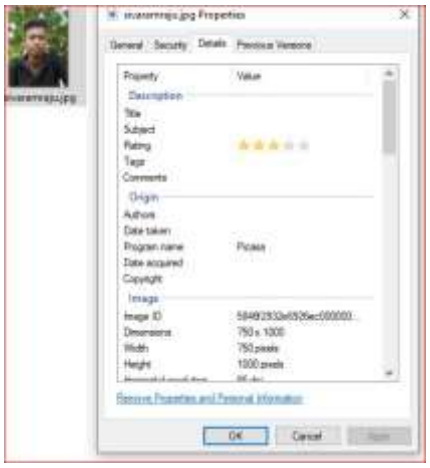
Search image2

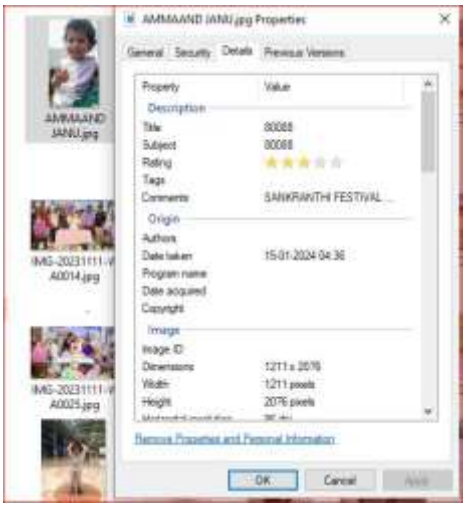
The Query image is not existing in the data set find top 1 similar image:-

If the query image Is not existing in the dataset the image is out of the data set. Then the image can find image based on the metadata we can merge the image vector and metadata vector (rating Tag) both similar rating images can find.

Note if the image contain any rating the metadata rating tag take default zero rating

Search single image:1





query image search image

Why This :-

If the query image does not exist in the dataset, the system retrieves similar images. However, this may result in irrelevant images being displayed. To avoid this, the system can filter similar images based on the rating tag. Additionally, users can modify the rating tag or use different metadata tags to refine the search results.

🔻 Features

1. ResNet-18 for Feature Extraction

Uses ResNet-18, a pre-trained deep learning model, to extract high-level feature representations from images.

The final classification layer is removed to obtain feature vectors instead of labels.

Images are processed using predefined transformations:

Resized to 224×224 pixels

Converted to a tensor

Normalized with mean = [0.485, 0.456, 0.406] and std = [0.229, 0.224, 0.225]

Extracted feature vectors are normalized for better similarity comparison.

2. FAISS HNSW Indexing for Fast Search

FAISS (Facebook AI Similarity Search) is used to perform efficient nearest-neighbor searches.

Uses Hierarchical Navigable Small World (HNSW) graphs for faster approximate search.

Instead of a brute-force search, FAISS stores and indexes feature vectors for fast retrieval.

3. PCA (Principal Component Analysis) for Dimensionality Reduction

Extracted features from ResNet are 512-dimensional.

PCA (Principal Component Analysis) reduces the dimensions to 128 while maintaining relevant information.

Reduces memory footprint and speeds up similarity searches.

4. Metadata Extraction (e.g., Image Ratings)

Reads EXIF metadata from images to extract rating information.

If a rating is available, it's converted into a vector and added to the index.

This allows more accurate searches by incorporating additional metadata.

5. Interactive Streamlit UI for Image Search

A Streamlit-based user interface allows:

Uploading multiple images for indexing. Using Ctrl+A

Uploading a query image to find similar images.

Specifying how many similar images (k) to retrieve.

Displaying results in an easy-to-view format.