Project 3 report

Name: Gautam Ajey Khanapuri  
Email: [Khanapuri.g@northeastern.edu](mailto:Khanapuri.g@northeastern.edu)  
Date: 10 February 2026

## Short Description of Overall project

In this project I have implemented a **CBIR system** that finds visually similar images in a database by analyzing color, texture, and spatial structure rather than relying on text metadata.

As suggested in the assignment description the system uses a **two-program architecture** for efficiency. **Program 1 (P1)** pre-processes the image directory and extracts feature vectors into CSV files. This computationally expensive step runs once per configuration. **Program 2 (P2)** performs fast matching by loading pre-computed features from the above csv files and compares them against a target image using various distance metrics.

The implementation supports multiple **feature types** including color histograms (RG chromaticity, RGB, Hue-Saturation, intensity), texture analysis (GLCM co-occurrence statistics, Laws filter responses), and edge-based features (Sobel magnitude and orientation). Images can be analyzed as a whole or divided into spatial regions (top/bottom, left/right, quadrants, center) to capture layout information.

Multiple **distance metrics** are implemented including histogram intersection, sum of squared differences, chi-squared distance, correlation, and cosine similarity. P2 combines these metrics using weighted averaging for flexible multi-feature matching.

For the **custom design**, the system was specialized for building retrieval, emphasizing texture features (Laws filters, GLCM) to distinguish brick, stone, and concrete materials, with supporting color and geometric information.

## Task 1: Baseline Matching

Implements baseline image matching using a 7×7 pixel square extracted from the center of each image as the feature vector. Similarity is measured using sum-of-squared-difference (SSD) between feature vectors, where lower values indicate greater similarity. This simple approach establishes a performance baseline for comparing more sophisticated methods.

A red rubber duck on a bench

AI-generated content may be incorrect.A bouquet of red roses in a white wrapper

AI-generated content may be incorrect.

Figure - 0968 Figure - 0547

A green and white striped shirt

AI-generated content may be incorrect.A red toy on a white plastic chair

AI-generated content may be incorrect.

Figure – 641 Figure 4 - 1013

1. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.1016.jpg" --- --- 0

2. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0986.jpg" --- --- 95.5714

3. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0641.jpg" --- --- 148

4. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0547.jpg" --- --- 338.116

5. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.1013.jpg" --- --- 350.605

## Task 2: Histogram matching

Implements color-based image matching using a 2D RG chromaticity histogram as the feature vector. For each pixel, the chromaticity values r = R/(R+G+B) and g = G/(R+G+B) are computed and mapped to a 16×16 bin histogram. The histogram is normalized by dividing each bin by the total pixel count to obtain probability distributions. Histogram intersection is calculated as the sum of minimum values across corresponding bins: intersection = Σmin(H₁[i], H₂[i]). Distance is computed as 1 - intersection, where lower values indicate greater similarity. The implementation uses custom code for both histogram computation and intersection calculation without relying on OpenCV's histogram functions.

Results for the image pic.0164.jpg are as follows. The difference in the professor’s results and mine is the presence of pic.489.jpg. I implemented rg 2d histogram using the 16 bins for each r and g and using intersection as the difference method. It is crucial to note that the second image in the professor’s result also has no red in it similar to pic.489.jpg. For me 1032.jpg came in 16th position with only minor difference in distance value.

*A computer screen with a login page

AI-generated content may be incorrect.*

Figure - 0080

*A squirrel running on a road

AI-generated content may be incorrect.*

Figure - 0489

*A red toy car on a pile of clothes

AI-generated content may be incorrect.*

Figure - 0461

1. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0164.jpg" --- --- 1.039161361e-07

2. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0080.jpg" --- --- 0.3085571081

3. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0489.jpg" --- --- 0.3447448488

4. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0461.jpg" --- --- 0.3560302921

5. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0898.jpg" --- --- 0.40047910455.

…

…

16. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0898.jpg" --- --- 0.46047810318

I ran the a Whole image RGB histogram and got exactly matching results with the professor. I used 8 bins for each channel and using intersection as the method to measure similarity.

A satellite dish on top of a building

AI-generated content may be incorrect.

Figure - 0110

Low angle view of a building

AI-generated content may be incorrect.

Figure - 1032

A stone building with a clock on the side

AI-generated content may be incorrect.

Figure - 0092

1. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0164.jpg" --- --- 9.040786608e-08

2. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0110.jpg" --- --- 0.5753692656

3. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.1032.jpg" --- --- 0.5765563892

4. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0092.jpg" --- --- 0.610900877

5. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0976.jpg" --- --- 0.6266387922

## Task 3: Multi Histogram Matching

Implements spatial image matching using multiple histograms from different image regions combined through weighted distance averaging. The system supports flexible spatial partitioning including top/bottom halves, left/right halves, quadrants, and center/edge regions. Each region can use different histogram types (RG chromaticity, RGB, Hue-Saturation) and distance metrics (intersection, chi-squared, SSD, correlation). The configuration string encodes part-histogram-metric-weight combinations, where weights are automatically normalized to sum to 1.0. For example, a configuration might use top-half RG chromaticity with intersection (weight 0.6) combined with bottom-half RGB using chi-squared (weight 0.4). Final distance is computed as the weighted sum of individual part distances: D\_final = Σ(w\_i × D\_i), where w\_i are normalized weights and D\_i are part-specific distances.

I ran the same specifications as described in the assignment and got nearly similar results. Images pic.0273.jpg and pic.1031.jpg appearing in the first six values as you can see in the output displayed in the box.

A large white water tower

AI-generated content may be incorrect.

Figure - 209

A large stone house with trees in front

AI-generated content may be incorrect.

Figure - 0409

A lamp post next to a building

AI-generated content may be incorrect.

Figure - 1055

1. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0274.jpg" --- --- 0.1023071828

2. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0209.jpg" --- --- 0.3271026505

3. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0409.jpg" --- --- 0.3504394529

4. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.1055.jpg" --- --- 0.3644256623

5. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.1031.jpg" --- --- 0.3811035189

Want to see the next image? Press 'n'.

If you want to exit, press 'q'.

n

6. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0273.jpg" --- --- 0.3814971849

## Task 6: Texture and Color

I have modelled the experiments for this task as a matrix. I have performed 8 queries. I have used intersection for color histograms and corelation for texture histograms. Target image for all will be pic.0535.jpg

A group of people sitting at a table

AI-generated content may be incorrect.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **2D rg Histogram**  16x16 bins | 3D RGB Hist (top and bottom separate)  8x8x8 bins |  |
| **1D Gradient Magnitude Hist**  1x16 | Experiment 5 | Experiment 7 | Experiment 3 |
| **2D Gradient Magnitude vs Gradient orientation Hist**  9x16 | Experiment 6 | Experiment 8 | Experiment 4 |
|  | Experiment 1 | Experiment 2 |  |

Implements combined texture and color analysis using multiple feature extraction methods. **Color features** include whole-image RG chromaticity histograms and spatially-split RGB histograms. **Texture features** use two gradient-based approaches: (1) 1D histogram of Sobel gradient magnitudes capturing edge strength distribution, and (2) 2D joint histogram of magnitude versus orientation (9 angle bins, 0-180°) capturing both edge strength and directional patterns. Distance metrics are matched to feature types: histogram intersection for color distributions and correlation for texture pattern matching. Eight experiments test different feature combinations with equal weighting to evaluate the contribution of spatial information and texture complexity to retrieval performance.

### Experiment 1

The images have been placed from left to right. The first image being most similar or at number 2 followed by the next images in the terminal output provided below.

1. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0535.jpg" --- --- 1.67840426e-07

2. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0733.jpg" --- --- 0.1215301343

3. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0731.jpg" --- --- 0.1344268605

4. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0340.jpg" --- --- 0.1374633415

5. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0003.jpg" --- --- 0.1387756097

A construction vehicle with a large bucket

AI-generated content may be incorrect.A yellow construction vehicle on gravel

AI-generated content may be incorrect.A person sitting on a chair

AI-generated content may be incorrect.

### Experiment 2

1. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0535.jpg" --- --- 1.456774044e-07

2. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0285.jpg" --- --- 0.277175893

3. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0628.jpg" --- --- 0.3267700204

4. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0952.jpg" --- --- 0.329403719

5. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0698.jpg" --- --- 0.3315002447

A long hallway with glass windows

AI-generated content may be incorrect.A room with snowflakes and balloons

AI-generated content may be incorrect.A stairs in a park

AI-generated content may be incorrect.

### Experiment 3

1. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0535.jpg" --- --- 4.440892099e-16

2. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0106.jpg" --- --- 6.764356036e-05

3. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0262.jpg" --- --- 7.841387451e-05

4. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0708.jpg" --- --- 0.000108449299

5. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0755.jpg" --- --- 0.0001279817288

A green dumpster on the side of a road

AI-generated content may be incorrect.A building with grass and trees

AI-generated content may be incorrect.A golf cart parked in front of a building

AI-generated content may be incorrect.

### Experiment 4

1. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0535.jpg" --- --- 1.698641228e-14

2. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0751.jpg" --- --- 0.01055499118

3. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0750.jpg" --- --- 0.01345809591

4. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0105.jpg" --- --- 0.01423613397

5. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0212.jpg" --- --- 0.01452239418

A green dumpster outside

AI-generated content may be incorrect.A green dumpsters next to a tree

AI-generated content may be incorrect.A yellow bulldozer on the street

AI-generated content may be incorrect.

### Experiment 5

1. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0535.jpg" --- --- 8.392021322e-08

2. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0733.jpg" --- --- 0.06186746616

3. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.1106.jpg" --- --- 0.08559785581

4. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0745.jpg" --- --- 0.08799111454

5. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0340.jpg" --- --- 0.08853490286

A construction vehicle with a large bucket

AI-generated content may be incorrect.A road sign on the side of a road

AI-generated content may be incorrect.A large orange excavator in front of a building

AI-generated content may be incorrect.

### Experiment 6

1. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0535.jpg" --- --- 8.392022149e-08

2. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0340.jpg" --- --- 0.09271085195

3. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0745.jpg" --- --- 0.09563181972

4. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.1106.jpg" --- --- 0.09725195589

5. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0339.jpg" --- --- 0.09789375885

A person sitting on a chair

AI-generated content may be incorrect.A large orange excavator in front of a building

AI-generated content may be incorrect.A road sign on the side of a road

AI-generated content may be incorrect.

### Experiment 7

1. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0535.jpg" --- --- 7.28387024e-08

2. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0285.jpg" --- --- 0.151534198

3. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0628.jpg" --- --- 0.1695818373

4. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0698.jpg" --- --- 0.1705793516

5. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0171.jpg" --- --- 0.1751790581

A long hallway with glass windows

AI-generated content may be incorrect.A room with snowflakes and balloons

AI-generated content may be incorrect.A white building with pillars

AI-generated content may be incorrect.

### Experiment 8

1. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0535.jpg" --- --- 7.283871067e-08

2. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0285.jpg" --- --- 0.1652887318

3. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0628.jpg" --- --- 0.1789458488

4. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0171.jpg" --- --- 0.1803817583

5. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0355.jpg" --- --- 0.1845573786

A long hallway with glass windows

AI-generated content may be incorrect.A room with snowflakes and balloons

AI-generated content may be incorrect.A cluttered room with a desk and a lamp

AI-generated content may be incorrect.

Results Analysis:   
Color-only matching (Exp 1, 2) retrieves images with similar color palettes regardless of scene content. The stone wall's brown/grey tones dominate matching. Texture-only matching (Exp 3, 4) retrieves images with similar structural patterns - edge density and orientation - completely ignoring color. This produces different matches that share geometric properties. Combined matching (Exp 5-8) balances both aspects. The 2D magnitude-orientation histogram (Exp 6, 8) provides more discriminative texture features than 1D magnitude alone (Exp 5, 7), as seen in the ranking differences. Spatial splitting (Exp 2, 7, 8) captures layout information but may reduce discrimination when scenes have uniform spatial structure.

## Task 5: Deep Network embeddings

Using the pre-computed 512-dimensional feature vectors extracted from the global average pooling layer of a ResNet18 network trained on ImageNet. Unlike classic features computed dynamically, DNN embeddings are loaded from a CSV database where each row contains an image filename followed by 512 floating-point values. Cosine similarity is used as the distance metric, calculated by normalizing both vectors by their L2-norm and computing d(v₁, v₂) = 1 - (v₁·v₂)/(||v₁|| × ||v₂||), where lower values indicate semantic similarity. The DNN features capture high-level semantic content learned from ImageNet's diverse object categories, providing complementary information to the hand-crafted color and texture features from previous tasks.

Cosine similarity on DNN embeddings for pic.0893.jpg.

A fire hydrant in the grass

AI-generated content may be incorrect.

1. "pic.0893.jpg" --- --- -2.220446049e-16

2. "pic.0897.jpg" --- --- 0.1517683092

3. "pic.0136.jpg" --- --- 0.1761571139

4. "pic.0146.jpg" --- --- 0.2248572818

5. "pic.0135.jpg" --- --- 0.2251176467

A yellow fire hydrant in the grass

AI-generated content may be incorrect.A silver fire hydrant in the grass

AI-generated content may be incorrect.A silver fire hydrant in the grass

AI-generated content may be incorrect.

Intersection on whole image RGB histogram for pic.0893.jpg:

1. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0893.jpg" --- --- 1.401031113e-07

2. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0136.jpg" --- --- 0.2659332435

3. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0897.jpg" --- --- 0.3077087449

4. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0368.jpg" --- --- 0.3327972361

5. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0123.jpg" --- --- 0.3472625885

A silver fire hydrant in the grass

AI-generated content may be incorrect.A yellow fire hydrant in the grass

AI-generated content may be incorrect.A book on the grass

AI-generated content may be incorrect.

Cosine similarity on DNN embeddings for pic.0164.jpg

A flag on top of a building

AI-generated content may be incorrect.

1. "pic.0164.jpg" --- --- 0

2. "pic.1032.jpg" --- --- 0.2121893879

3. "pic.0213.jpg" --- --- 0.2128362677

4. "pic.0690.jpg" --- --- 0.2351369888

5. "pic.0426.jpg" --- --- 0.2492830327

Low angle view of a building

AI-generated content may be incorrect.A building with trees around it

AI-generated content may be incorrect.A tall stone tower with a few windows

AI-generated content may be incorrect.

Intersection on whole image RGB histogram for pic.0164.jpg:

1. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0164.jpg" --- --- 1.107628123e-07

2. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0110.jpg" --- --- 0.6147949368

3. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.1032.jpg" --- --- 0.6764526393

4. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0092.jpg" --- --- 0.6940948393

5. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0976.jpg" --- --- 0.7220153794

A satellite dish on top of a building

AI-generated content may be incorrect.Low angle view of a building

AI-generated content may be incorrect.A stone building with a clock on the side

AI-generated content may be incorrect.

#### Result Analysis:

For pic.0893.jpg – The first two images are identical to the results of the DNN query. However, the third image is of a text book on the grass. On closer observation it is fairly obvious why RGB histograms match. The background is green in both cases. The green part of the book is also there in the target fire hydrant. The shine on the fire hydrant which is white, grey and silver can easily match with the part of the book shining and the person’s shoe visible in the image. Since there is no spatial discrimination, the RGB metric places them quite close by purely on the similarity in colour. But the DNN embeddings have obviously been trained. All fire hydrants would have been labelled as fire hydrant and all books would have been labelled as books. Therefore it is unlikely that these two would have similar vectors and as result are placed far apart by the DNN.

For pic.0164.jpg – If we notice the color in the images produced by the classic method: it is clear. The sky has the same colour in three images. But in the DNN images, the colour of the sky is different. There are more clouds. But most likely those images must have all been labelled as a ‘tower’ regardless of the colour of the sky. That’s why there is a difference between DNN’s output and classic method’s output.

## Task 6: Compare DNN Embeddings and Classic Features

Here I evaluate whether deep network embeddings or hand-crafted features produce superior retrieval results across different query types. But "better" is defined contextually based on user intent: semantic similarity (retrieving same object category regardless of appearance) versus visual similarity (retrieving similar colors, textures, and spatial layouts). DNN embeddings excel at semantic matching due to ImageNet training on object categories, while classic features excel at perceptual matching based on low-level visual properties. Take a look at the images that I have examined and analysed:

Query Picture: pic.1072.jpg

A close-up of a bush of pink roses

AI-generated content may be incorrect.

DNN results:

1. "pic.1072.jpg" --- --- 0

2. "pic.0143.jpg" --- --- 0.1610317034

3. "pic.0863.jpg" --- --- 0.2004461723

4. "pic.0329.jpg" --- --- 0.2071869466

5. "pic.0144.jpg" --- --- 0.2074501887

A close-up of a garden

AI-generated content may be incorrect.Purple flowers in the ground

AI-generated content may be incorrect.A close-up of a red flower

AI-generated content may be incorrect.

Classic results:

1. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.1072.jpg" --- --- 2.508817811e-07

2. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0701.jpg" --- --- 0.3049797788

3. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0813.jpg" --- --- 0.3075357018

4. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.1069.jpg" --- --- 0.3224522694

5. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0899.jpg" --- --- 0.3294896411

A street light in a park

AI-generated content may be incorrect.A wooden bench in a forest

AI-generated content may be incorrect.A group of rocks and leaves

AI-generated content may be incorrect.

Analysis:

The images show a slight difference. The Classic method I have used is RGB hist for top and bottom each weighted 3units and Sobel mag vs orientation histogram weighted 1unit. The classic method picks up images with a large amount of greenery and foliage. The red coloured pixels are very few and hence missing these pixels will not adversely affect the distance. Therefore we are seeing images with colour mostly matching. Since there is not a lot of regular pattern in the target image, the sobel histogram is not able to be of much service. The DNN embeddings are tuned to keep all images having flowers close to one another regardless of colour or background.

Query Picture: pic.0948.jpg

A stuffed animal on a staircase

AI-generated content may be incorrect.

DNN results:

1. "pic.0948.jpg" --- --- 1.110223025e-16

2. "pic.0930.jpg" --- --- 0.1283092643

3. "pic.0960.jpg" --- --- 0.2003666104

4. "pic.0928.jpg" --- --- 0.2040939423

5. "pic.0972.jpg" --- --- 0.2167987105

A small stuffed animal with a red shirt

AI-generated content may be incorrect.A stuffed animal holding a flower

AI-generated content may be incorrect.A stuffed animal in the grass

AI-generated content may be incorrect.

Classic results:

1. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0948.jpg" --- --- 1.021100112e-07

2. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0217.jpg" --- --- 0.2413966826

3. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0675.jpg" --- --- 0.2509134216

4. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0617.jpg" --- --- 0.2696716225

5. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0952.jpg" --- --- 0.2789157882

A wall of safes with numbers and numbers

AI-generated content may be incorrect.A window in a stone building

AI-generated content may be incorrect.A close-up of a stone building

AI-generated content may be incorrect.

Analysis:

The image is that of a bunny. The DNN embeddings have been trained in that manner. Therefore, the DNN embeddings successfully put the images of the stuffed bunny together despite the background colour distribution, brightness of the three images greatly varying. However, in the classic results notice that the brightness and colours of all 4 images are nearly identical. Except for the colour red, all other colours are present in the result images. Moreover the presence of stairs, produce strong strong gradient responses in the sobel filters. Thus the images chosen also have quite a few marked gradients.

Query Picture: pic.0734.jpg

A yellow roller on a construction site

AI-generated content may be incorrect.

DNN results:

1. "pic.0734.jpg" --- --- 0

2. "pic.0731.jpg" --- --- 0.1549314665

3. "pic.0735.jpg" --- --- 0.1653996604

4. "pic.0739.jpg" --- --- 0.1829307486

5. "pic.0743.jpg" --- --- 0.1872635715

A yellow construction vehicle on gravel

AI-generated content may be incorrect.A construction site with a white vehicle

AI-generated content may be incorrect.A construction vehicle digging a hole

AI-generated content may be incorrect.

Classic results:

1. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0734.jpg" --- --- 1.797431813e-07

2. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0001.jpg" --- --- 0.3179226137

3. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0577.jpg" --- --- 0.320950287

4. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0065.jpg" --- --- 0.3281675094

5. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0733.jpg" --- --- 0.333551052

A group of people in a room

AI-generated content may be incorrect.A bush with red leaves

AI-generated content may be incorrect.A window with trees in the background

AI-generated content may be incorrect.

Analysis:

Amongst the four images selected by the classic method, nothing semantically is matching. They are images of completely different things. But that is what we perceive. All these four images have the same colour distribution as the target image. The computer does not see faces. It only sees colours, gradients, patterns in the pixel values. On the other hand DNN embeddings have been trained to resemble human perception of similarity.

What is better:  
No single method is perfect. Both have it’s own strengths and weaknesses. But for this dataset, DNN is superior in finding images of other flowers when given an image of a flower. The choice depends on whether we want semantic similarity or visual similarity.

## Task 7: Custom Design

**Category Selection:** Building and architectural images. I have chosen this as the target category due to their distinctive visual characteristics and the opportunity to leverage texture analysis methods.

**Feature Design Rationale:**

Buildings are primarily distinguished by **material textures** (brick, stone, concrete, wood), **architectural geometry** (vertical lines, windows, structural elements), and **color properties** (material colors, sky conditions). The custom feature set emphasizes texture analysis:

**Bottom-half Laws filter histograms (weight 0.35)** - Captures material texture patterns. Laws filters detect edges, spots, and ripples characteristic of brick mortar, stone surfaces, and concrete finishes. Correlation metric matches texture patterns.

**Whole-image Sobel magnitude-orientation histogram (weight 0.25)** - Captures architectural geometry. Buildings exhibit strong vertical and horizontal edges from walls, windows, and structural elements. Correlation metric identifies similar geometric patterns.

**Bottom-half RGB histogram (weight 0.20)** - Differentiates material colors: red brick, grey stone, white concrete. Chi-squared distance measures color distribution similarity.

**Top-half Hue-Saturation histogram (weight 0.15)** - Captures sky conditions and rooflines. Intersection metric for color matching.

**Design Philosophy:** Texture features receive 60% combined weight because material properties (brick vs. stone) are more distinctive for buildings than color alone. Spatial splitting separates architectural elements (bottom) from sky/background (top).

**Evaluation:** Two queries test retrieval of different building types (stone masonry vs. brick structures) with analysis of top 5 matches and least similar images to demonstrate discriminative power.

Query img 1 – pic.0212.jpg

A stone building with stairs

AI-generated content may be incorrect.

1. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0212.jpg" --- --- 1.738356105e-07

2. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0892.jpg" --- --- 0.1411938047

3. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0891.jpg" --- --- 0.164000218

4. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0992.jpg" --- --- 0.1905655588

5. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0542.jpg" --- --- 0.1919213014

Want to see the next image? Press 'n'.

If you want to exit, press 'q'.

n

6. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0104.jpg" --- --- 0.1972052278

Want to see the next image? Press 'n'.

If you want to exit, press 'q'.

n

7. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0576.jpg" --- --- 0.1983972382

Want to see the next image? Press 'n'.

If you want to exit, press 'q'.

n

8. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0160.jpg" --- --- 0.1984885056

Want to see the next image? Press 'n'.

If you want to exit, press 'q'.

n

9. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0998.jpg" --- --- 0.2034026278

Want to see the next image? Press 'n'.

If you want to exit, press 'q'.

n

10. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0590.jpg" --- --- 0.2051799896

Want to see the next image? Press 'n'.

A lamp post in front of a stone building

AI-generated content may be incorrect.A lamp post next to a tree

AI-generated content may be incorrect.A close-up of a stone path

AI-generated content may be incorrect.A close-up of a wire

AI-generated content may be incorrect.A stone stairs in the woods

AI-generated content may be incorrect.

Analysis:  
The custom building retrieval system successfully identified stone and brick buildings (pic.0892, pic.0891, pic.0992) in the top 5 results, demonstrating that Laws texture features effectively discriminate material patterns.

However, some non-building images appear due to textural similarity. pic.0542 (cables on pavement) and pic.0590 (camera on concrete) ranked highly because rough pavement and concrete surfaces share similar texture responses to stone building materials. This illustrates a fundamental limitation of texture-only matching: materials can be similar across different object categories.

A black camera on a grey surface

AI-generated content may be incorrect.

This could be improved by: (1) adding spatial context (buildings typically have vertical structures), (2) incorporating shape features (rectangular windows, building silhouettes), or (3) increasing weight on geometric features (Sobel orientation) to emphasize architectural lines over ground textures.

Despite this the system is very successful in picking up blocks or stoned walls. Since I have heavily weighted Laws filters on the top half of the image, the sytem is picking up images even if there are building or walls in the background of the image. A human might miss them at first glance but there’s a building in the background of images. For example take this image pic.0555.jpg which is ranked 29th has a building in the background: Take another image pic.0562.jpg which is 51st. It too has buildings in the background. Since I have created a 3D RGB histogram on the bottom half of the image, images containing roads or pavements are getting included. Even the 100th image(pic.0155.jpg) is that of a tiled pavement and there is a stone wall in the background. The system is successfully picking up block structures even in the 150th image.  
A large green box on the side of a road

AI-generated content may be incorrect.A silver car parked in a parking lot

AI-generated content may be incorrect.A circular dart board on a bench

AI-generated content may be incorrect.

Query img 2 – pic.0632.jpg

A stone house with snow

AI-generated content may be incorrect.

1. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0632.jpg" --- --- 1.546778861e-07

2. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.1101.jpg" --- --- 0.1092522795

3. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0971.jpg" --- --- 0.1624056009

4. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0615.jpg" --- --- 0.1843239925

5. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0041.jpg" --- --- 0.188221964

Want to see the next image? Press 'n'.

If you want to exit, press 'q'.

n

6. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0569.jpg" --- --- 0.1882312366

Want to see the next image? Press 'n'.

If you want to exit, press 'q'.

n

7. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0658.jpg" --- --- 0.198087462

Want to see the next image? Press 'n'.

If you want to exit, press 'q'.

n

8. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0064.jpg" --- --- 0.199495379

Want to see the next image? Press 'n'.

If you want to exit, press 'q'.

n

9. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0996.jpg" --- --- 0.200211922

Want to see the next image? Press 'n'.

If you want to exit, press 'q'.

n

10. "/Users/ajeyk/neu/cvpr/proj\_2/image\_database/olympus/pic.0629.jpg" --- --- 0.2027585668

A building with a bell tower

AI-generated content may be incorrect.A close-up of a stained glass window

AI-generated content may be incorrect.A tree with a group of animals in it

AI-generated content may be incorrect.A person wearing sunglasses and looking at a computer screen

AI-generated content may be incorrect.A dome on a building

AI-generated content may be incorrect.

Analysis:

I deliberately picked a more difficult image this time. The image has no major brick pattern in the lower half where much of the laws text matching happening. Despite this it has a success rate of 3 out of 5 in picking up images of building.   
In particular the fourth image is intriguing. it appears to be the back of a person’s head wearing sun glasses.

This match likely occurred due to: (1) vertical architectural elements (walls, frames) triggering similar Sobel orientation patterns, (2) neutral color palette (greys, browns) matching the stone tones in the query's lower half, and (3) smooth surface textures that correlate with the non-brick portions of the query building.

This demonstrates both a strength and limitation: the system successfully generalizes beyond brick-specific textures (matching stone, painted surfaces, and architectural elements), but occasionally due to the lack of structure in the bottom half of the image, it is matching with similar images that donot have boxed structure in the bottom half. Or it could be the white matching with all the snow in the image. In order to improve performance on this particular image, I would add another laws texture feature for the center and not just the bottom.

## Task 7: Extension

Extension 1: Advanced Texture Analysis Methods

Implemented two sophisticated texture analysis approaches beyond basic Sobel gradients:

**GLCM (Gray-Level Co-occurrence Matrix) Features:** Extracts five statistical texture measures (energy, contrast, homogeneity, entropy, maximum probability) from co-occurrence matrices computed at four spatial offsets [(1,0), (0,1), (1,1), (2,0)]. This captures texture at multiple scales and orientations, producing 20-dimensional feature vectors that characterize spatial intensity relationships.

**Laws Filter Response Histograms:** Applies nine Laws filter combinations (L5E5, E5L5, E5E5, S5S5, L5S5, S5L5, W5W5, R5R5, E5S5) to detect edges, spots, waves, and ripples. Each filter's response is histogrammed into 16 bins, creating 144-dimensional texture signatures. This multi-scale approach effectively discriminates between brick, stone, and concrete materials.

**Impact:** Texture features significantly improved building retrieval accuracy, enabling material-based discrimination that color histograms alone cannot achieve.

**Extension 2: Comprehensive Distance Metric Suite**

Implemented ten distance metrics including standard measures (SSD, histogram intersection, chi-squared, cosine similarity, correlation, Bhattacharyya, Manhattan, Earth Mover's Distance) and two custom metrics:

**Custom Perceptual Distance:** Weights histogram bins based on perceptual importance, emphasizing mid-range values over extremes to match human visual sensitivity.

**Custom Weighted Combination:** Application-specific metric designed for the chosen image category, combining multiple standard metrics with domain knowledge.

**Impact:** Enables optimal metric selection per feature type (intersection for color, correlation for texture patterns, cosine for embeddings).

**Extension 3: Flexible Multi-Feature Architecture**

Designed an extensible configuration system supporting arbitrary combinations of spatial regions, histogram types, distance metrics, and weights through compact encoding strings. The modular design uses function registries and enums for dynamic dispatch, allowing new features or metrics to be added without modifying core matching logic.

**Impact:** Enables rapid experimentation with feature combinations and supports complex weighted multi-part matching strategies.

I have used these features in my custom CBIR implementation:  
./p2 image\_database/olympus/pic.0632.jpg -m-TGO7wSO5TRq4thI3 image\_database/olympus/bottom\_law\_multi\_histogram\_ft\_vec\_1770880794.csv image\_database/olympus/whole\_sobel\_magnitude\_vs\_orientation\_2d\_multi\_histogram\_ft\_vec\_1770880796.csv image\_database/olympus/bottom\_rgb\_multi\_histogram\_ft\_vec\_1770880834.csv image\_database/olympus/top\_hs\_multi\_histogram\_ft\_vec\_1770880837.csv

**Configuration:** -m-TGO7wSO5TRq4thI3

This encodes a custom 4-part weighted feature combination for building retrieval:

**Part 1: TGO7**

T = Bottom half

G = Laws filter response histograms (144 bins: 9 filters × 16 bins)

O = Correlation distance metric

7 = Weight (36.8% after normalization)

**Purpose:** Captures material texture patterns (brick, stone, concrete)

**Part 2: wSO5**

w = Whole image

S = Sobel magnitude vs. orientation 2D histogram (144 bins: 16 mag × 9 angle)

O = Correlation distance metric

5 = Weight (26.3%)

**Purpose:** Captures architectural geometry (vertical/horizontal lines, structural patterns)

**Part 3: TRq4**

T = Bottom half

R = RGB color histogram (32,768 bins: 32×32×32, flattened)

q = Chi-squared distance metric

4 = Weight (21.1%)

**Purpose:** Distinguishes material colors (red brick, grey stone, white concrete)

**Part 4: thI3**

t = Top half

h = Hue-Saturation histogram (1024 bins: 32×32)

I = Intersection distance metric

3 = Weight (15.8%)

**Purpose:** Matches sky conditions and upper architectural elements

**Final distance:** Weighted combination emphasizing texture (63.1% combined) over color (36.9%), optimized for material-based building discrimination.

# Reflection

This project deepened my understanding of content-based image retrieval beyond simple pixel comparisons. I learned that image similarity is multifaceted - color histograms capture appearance, texture features (GLCM, Laws filters) capture material properties, and spatial layouts capture compositional structure. Each captures different aspects of "similarity," and there's no single "correct" answer.

The two-program architecture taught me the value of preprocessing for computational efficiency - separating feature extraction from matching enabled rapid experimentation with distance metrics and weights. Designing the configuration system with enums and function registries showed me how to build extensible, maintainable code rather than rigid if-else chains.

Debugging the Rect(x,y) vs (row,col) confusion and reference member initialization issues reinforced the importance of understanding C++ fundamentals deeply. Implementing texture analysis methods like GLCM and Laws filters gave me appreciation for how mathematical abstractions (co-occurrence statistics, filter banks) translate to practical discriminative power.

The custom CBIR task demonstrated that domain knowledge matters - understanding building characteristics (material textures, architectural geometry) guided effective feature design. The system's occasional "mistakes" (matching pavement textures) revealed fundamental limitations of appearance-based matching and the need for higher-level semantic understanding.

Overall, this project bridged the gap between low-level image processing and practical computer vision applications, showing how thoughtful feature engineering enables intelligent image understanding.

# Acknowledgement

I used the internet for a reading and understanding opencv’s documentation and use cases. I also used AI (Claude) to debug certain errors during installation, making of Makefile, usage of vim. I also used Claude for debugging my C++ code especially closer to the submission deadline. I also used it help me organize the functions properly into classes and make the program extendible. I used claude to generate comments for the code that I designed and wrote.