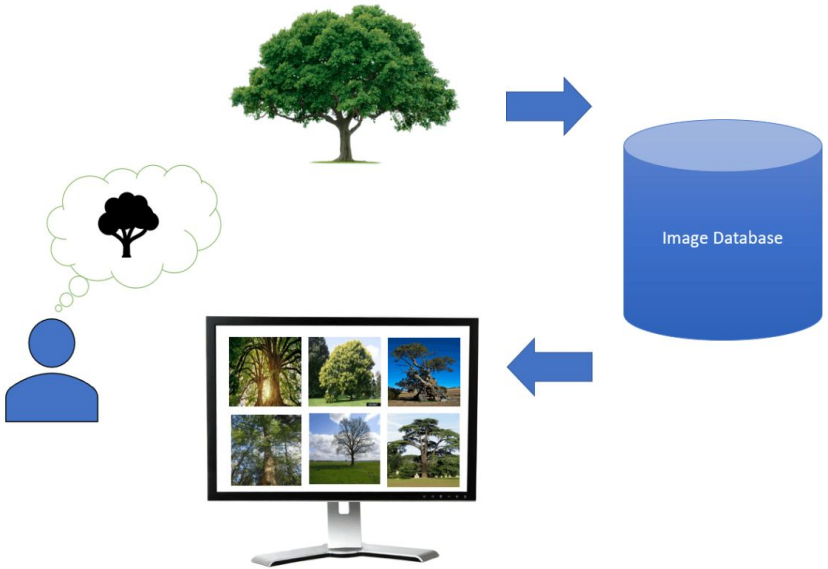
Image-to-image Retrieval

CS 3570 Multimedia

Introduction

- Description: Image-to-image retrieval is a task that aims to find images in a large database that are similar to a given query image



Related Work

- CLIP

```
Top 10 most similar images:
1. image.orig/298.jpg - Similarity Score: 0.67
2. image.orig/288.jpg - Similarity Score: 0.21
3. image.orig/287.jpg - Similarity Score: 0.07
4. image.orig/285.jpg - Similarity Score: 0.02
5. image.orig/289.jpg - Similarity Score: 0.02
6. image.orig/281.jpg - Similarity Score: 0.01
7. image.orig/284.jpg - Similarity Score: 0.00
8. image.orig/286.jpg - Similarity Score: 0.00
9. image.orig/283.jpg - Similarity Score: 0.00
10. image.orig/292.jpg - Similarity Score: 0.00
```

query img: 299.jpg

298.jpg

288.jpg

- Except for CLIP, You are still encouraged to try other methods

Dataset

- Our dataset is derived from parts of the Animal-10
- Total 10 category, including dog, cat, horse, elephant, butterfly, chicken, cow, sheep, squirrel, and spider
- 10000 images for training, and 1000 images for testing
- The divided dataset has been announced on the eeclass discussion board

Referece: https://kaggle.com/datasets/alessiocorrado99/animals10

Evaluation

- Precision and Recall
  - Precision = A / B
    - A: # of relevant retrieved images
    - B: # of total retrieved images
  - Recall = A / C
    - C: # of relevant images in database
- Speed
  - Measure the execution time
  - For fairness, TA will test your execution time on the same computer, but you are still encouraged to test it by yourselves

Rules

- Despite of pursuing performance, you are encouraged to provide some novelty and ingenuity
- You can’t train on test set
- No copying allowed

Reference

- A Comprehensive Analysis on Deep Learning based Image Retrieval
  - https://ieeexplore.ieee.org/document/10200622
- Learning Transferable Visual Models From Natural Language Supervision
  - https://arxiv.org/abs/2103.00020
- Awesome and classical image retrieval papers
  - https://github.com/willard-yuan/awesome-cbir-papers
- Image Retrieval on Real-life Images with Pre-trained Vision-and-Language Models
  - https://arxiv.org/abs/2108.04024