Large Scale Content-based Image Retrieval

Shaoting Zhang

Xiang Yu

RUTGERS
School of Arts and Sciences

Computational Biomedicine
Imaging & Modeling Center

Content-based Image retrieval

- Given an input image, find relevant / similar ones in the database.

- Use local and global image features.
- Large scale image retrieval: find similar images from millions of training images.

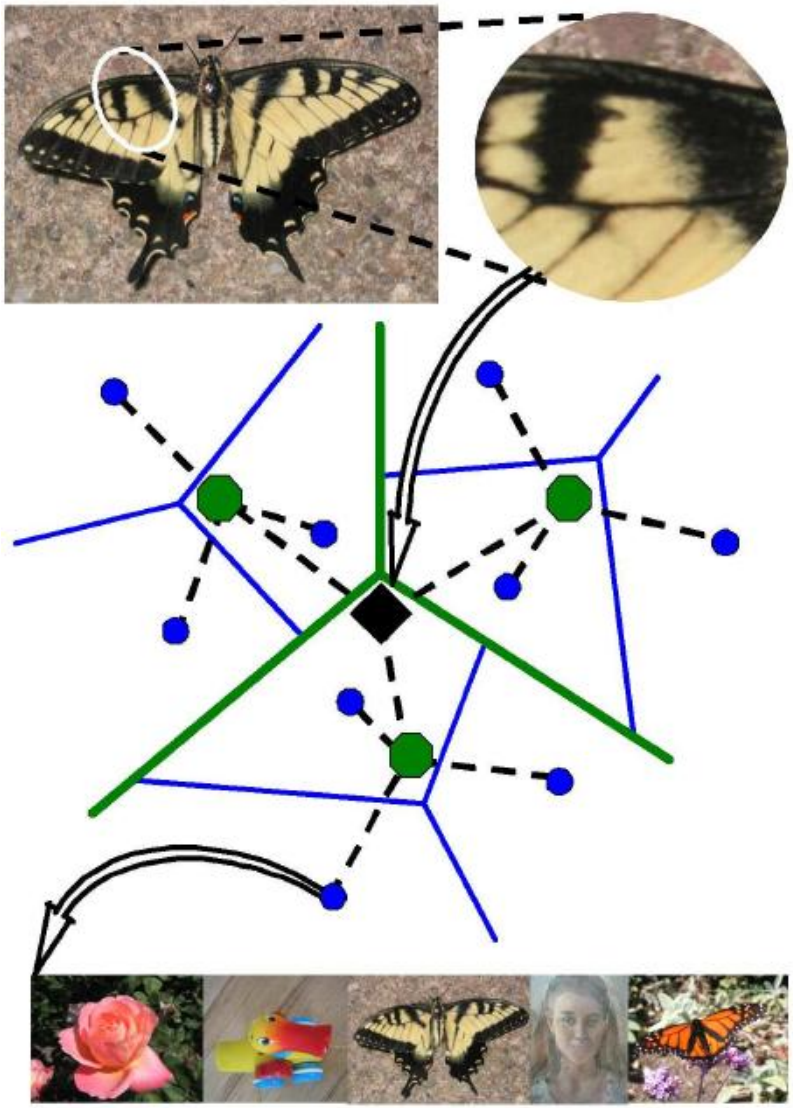
Outline

- **Efficient Local Feature**
  - Vocabulary Tree
  - City-Scale Landmark Identification
  - Results and Problems
- **Efficient Global Feature**
  - GIST, Color Features, Small Code
  - Corel 5K, UK bench, Results
- **Combination**
  - Motivation
  - Graph Fusion
  - Results and Discussions

Local Feature

Vocabulary Tree

- Extract descriptors (e.g., SIFT features).
- Hierarchical quantization instead of standard K-mean.
- Build inverted files with references to images containing an instance of that descriptor.
- Very efficient.

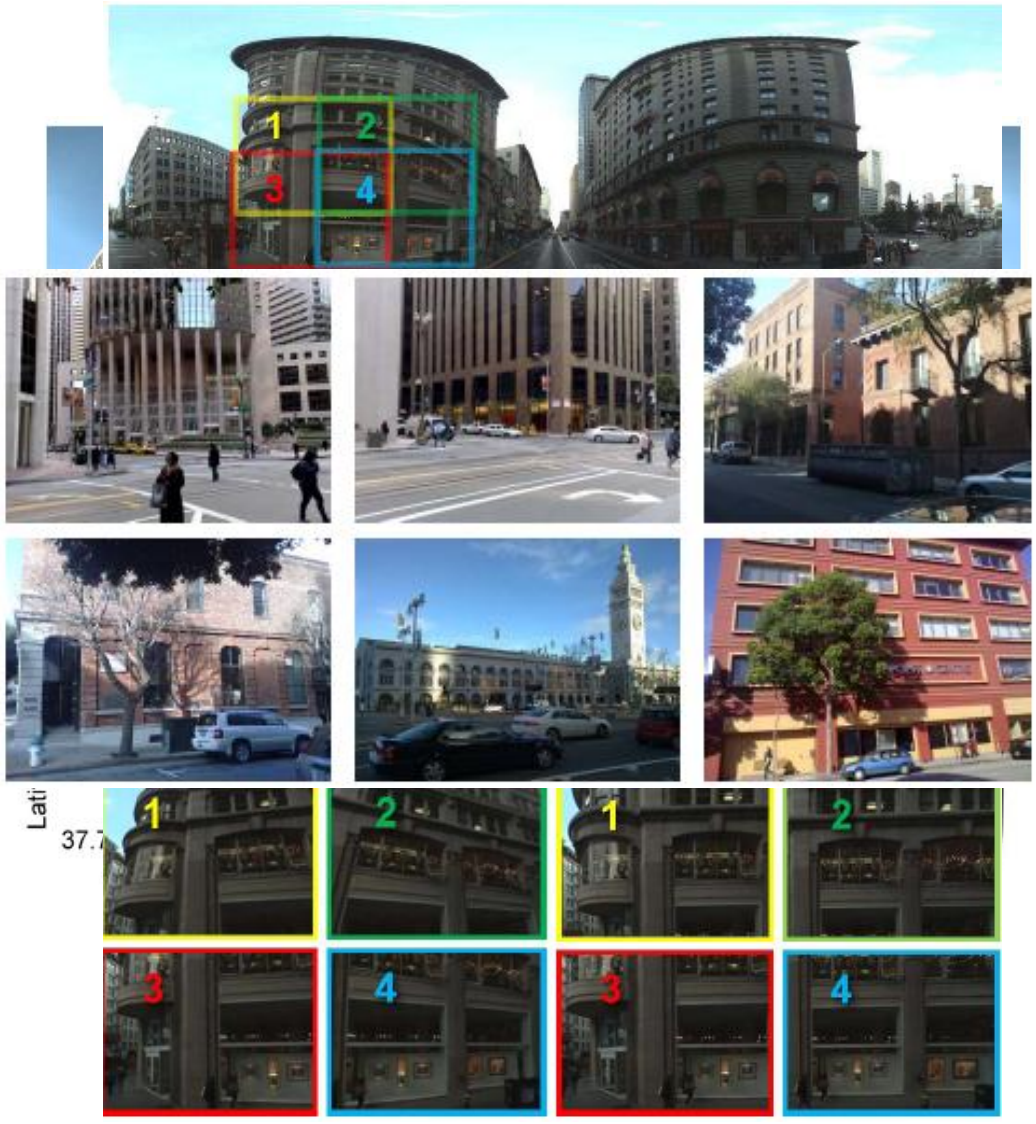


D. Nister and H. Stewenius, CVPR'06

Local Feature

City-Scale Landmark Identification

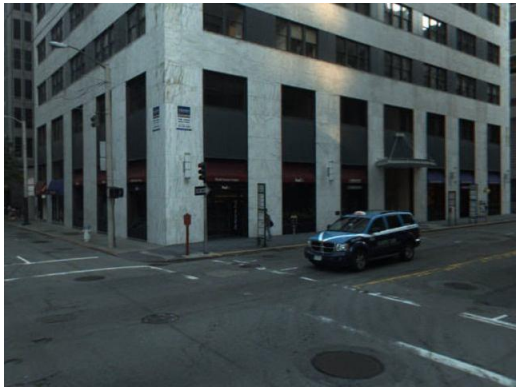
- Panorama images from San Francisco data.
- Application: query image taken with a smart phone. Then retrieve building image in database and its information.
- The largest set (1.7M).
- Perspective central and frontal images.
- Examples of query images



Devices, D. Chen, et.al., CVPR'11

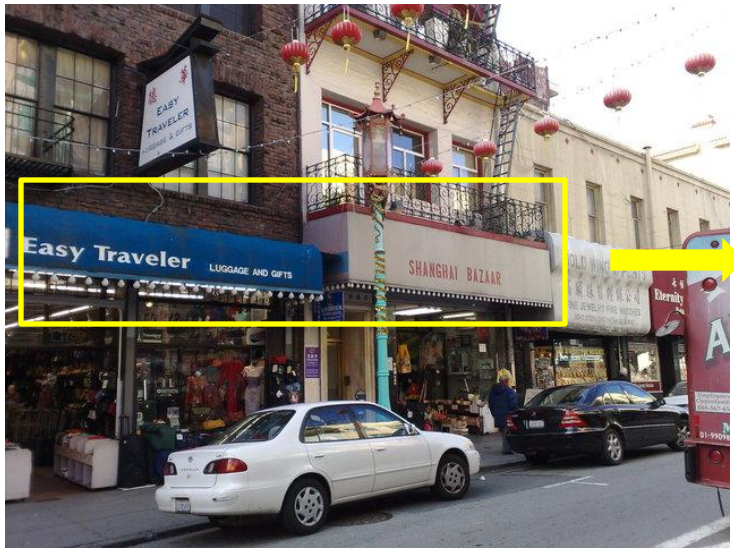
Local Feature

Our Retrieval Results



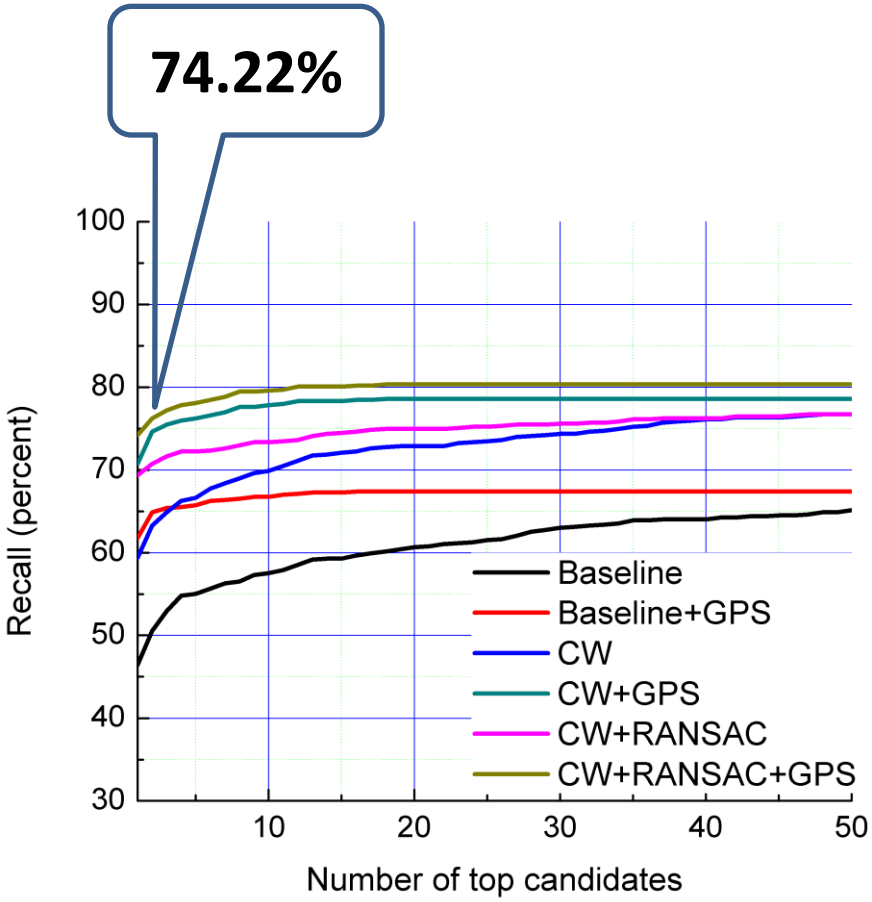
Local Feature

Our Retrieval Results

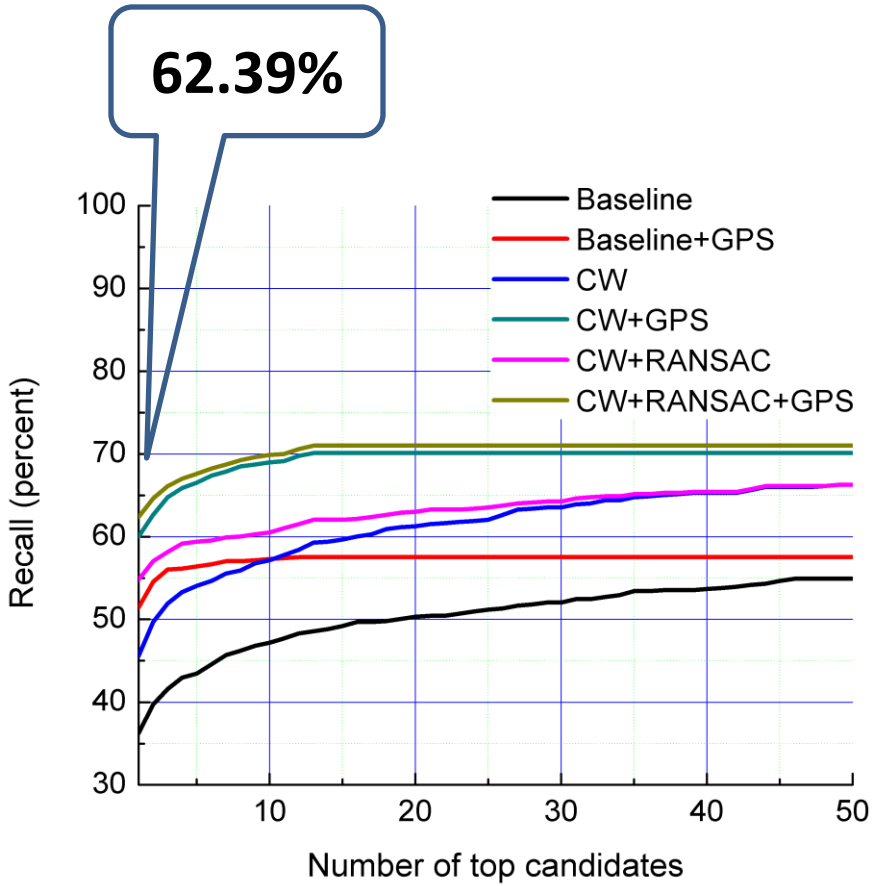


Local Feature

Our Retrieval Results



Perspective Central Images



Perspective Frontal Images

Local Feature

Problems of Local Features

- Local similarity may not generate correct results.
- Potential solution: Consider to use global features.

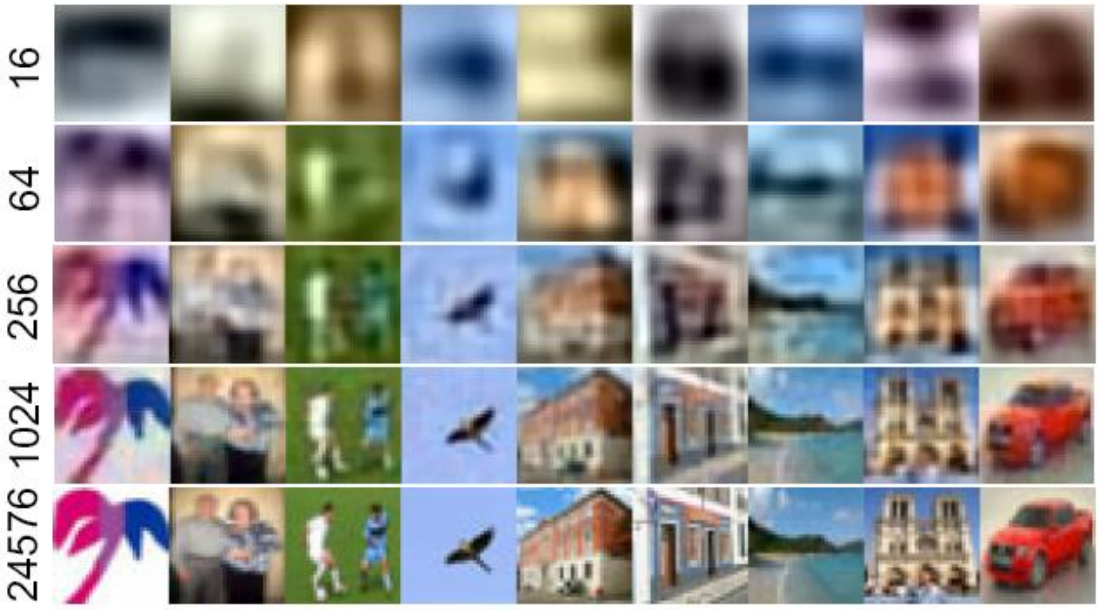
Query

Retrieval

Global Feature

Color, GIST, etc.

- RGB, LAB, HSV, 1D or 3D histogram.
- GIST (accumulating image statistics over the entire scene).
- Small code technique to accelerate the computation.



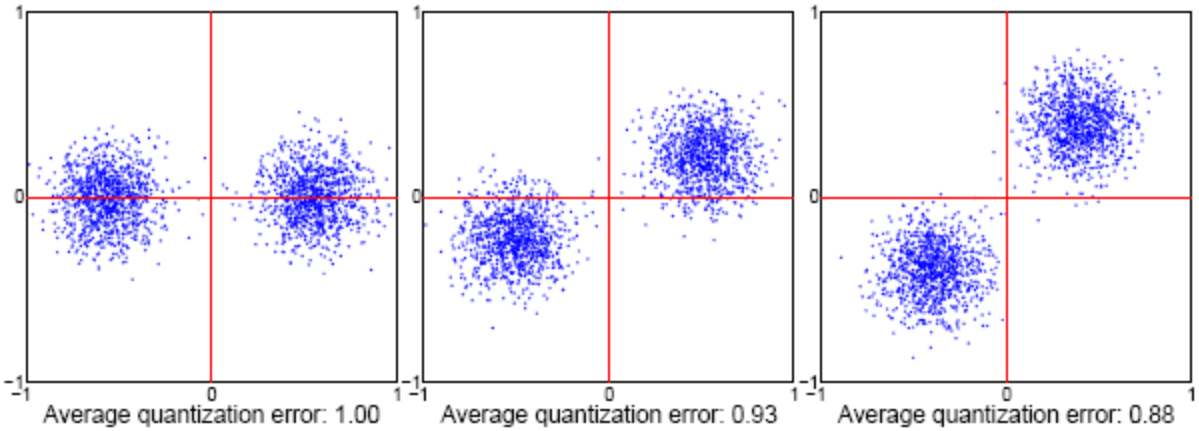
A. Oliva and A. Torralba, IJCV'01

A. Torralba, R. Fergus, Y. Weiss, CVPR'08

Global Feature

Small Code Technique

- PCA to reduce the dimension (960 bins -> 256 bins).
- Random rotation or optimized rotation.
- Binary quantization. Using Hamming distance.
- 960 floats -> 256 floats -> 256 bits (217 times smaller).



(a) PCA aligned. (b) Random Rotation. (c) Optimized Rotation.

Y. Gong and S. Lazebnik, CVPR'11

Global Feature

Corel 5k and UK bench

- Corel 5K: 50 categories, each category has 100 images. Leave-one-out for retrieval.
- Precision of Top-N retrievals.





Global Feature

Corel 5k and UK bench

- UK bench: 10200 images. 2550 objects. Each one has four images.
- Evaluation: 4 x recall at the first four returned images, referred as N-S score (maximum = 4).

D. Nister and H. Stewenius, CVPR'06

Global Feature

Results

- PCA tries to preserve L2 distance.
- GIST performs well using L2, while HSV prefers L1 or Bhattacharyya distance.
- Corel: we choose GIST. (VOC, 46.6%)
- UK: we choose HSV3D, NS = 3.17. (VOC, NS = 3.53)

Features	L1	L2	PCA	Binary	Random	ITQ
GIST	46.2%	**45.3%**	41.7%	33.1%	**42.5%**	40.6%
HSV	45.9%	**31.5%**	--	--	**34.8%**	--
HSV3D	54.3%	**35.8%**	--	--	**38.4%**	--

Corel 5k, Top-1 retrieval

Combination

Motivation

- Both global and local achieve good performance.

Dataset	Global	Local
Corel 5K	46.2%	46.6%
UK bench	3.17	3.53

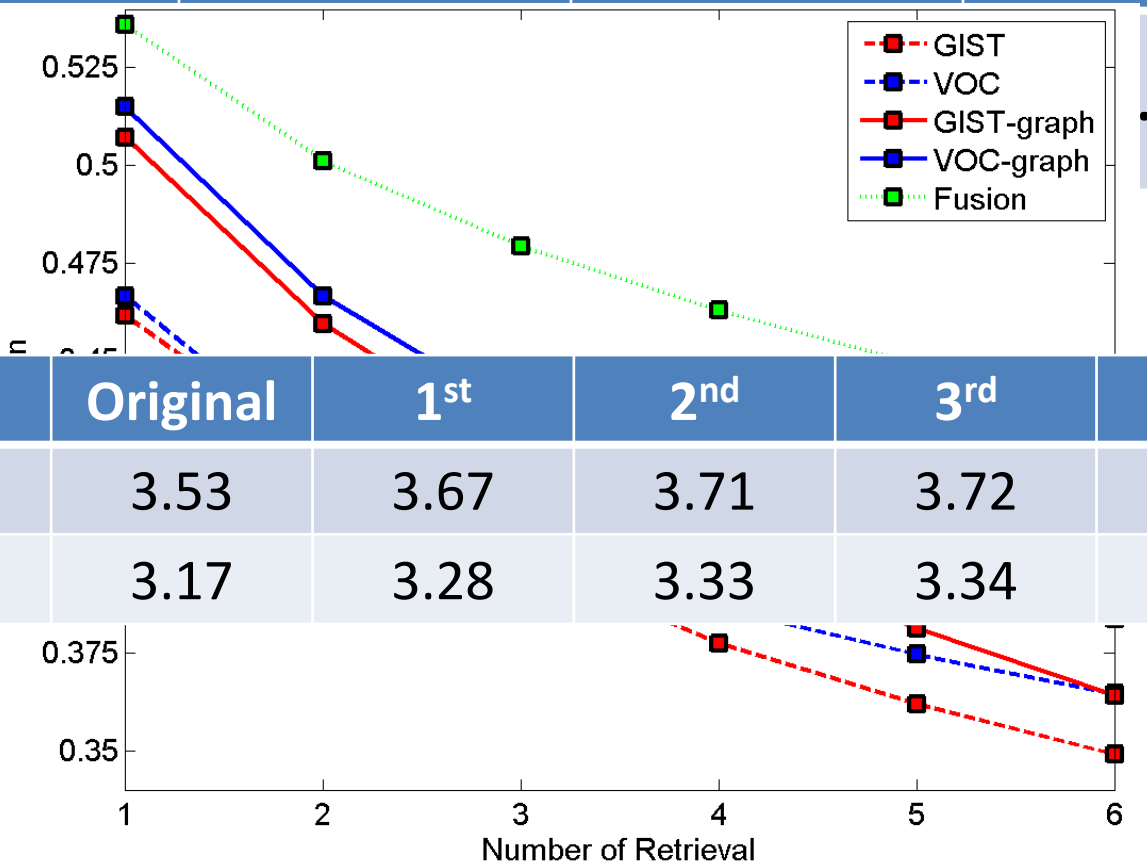
- When one fails, the other may do well. For top-5 retrievals of Corel 5K dataset:
  - Global feature fails to retrieve any correct images in **1,566** (out of 5,000) queries. In these 1,566 cases, local does well in **403** (>=2 correct).
  - Local feature fails in **1,671**, while global does well in **431**.

Combination

Results

- UK bench (NS score, state-of-the-art: 3.68):
- Corel 5K (Precision of top-N retrievals):

	Baseline	Single Graph	Graph Fusion
VO			.77
HS			

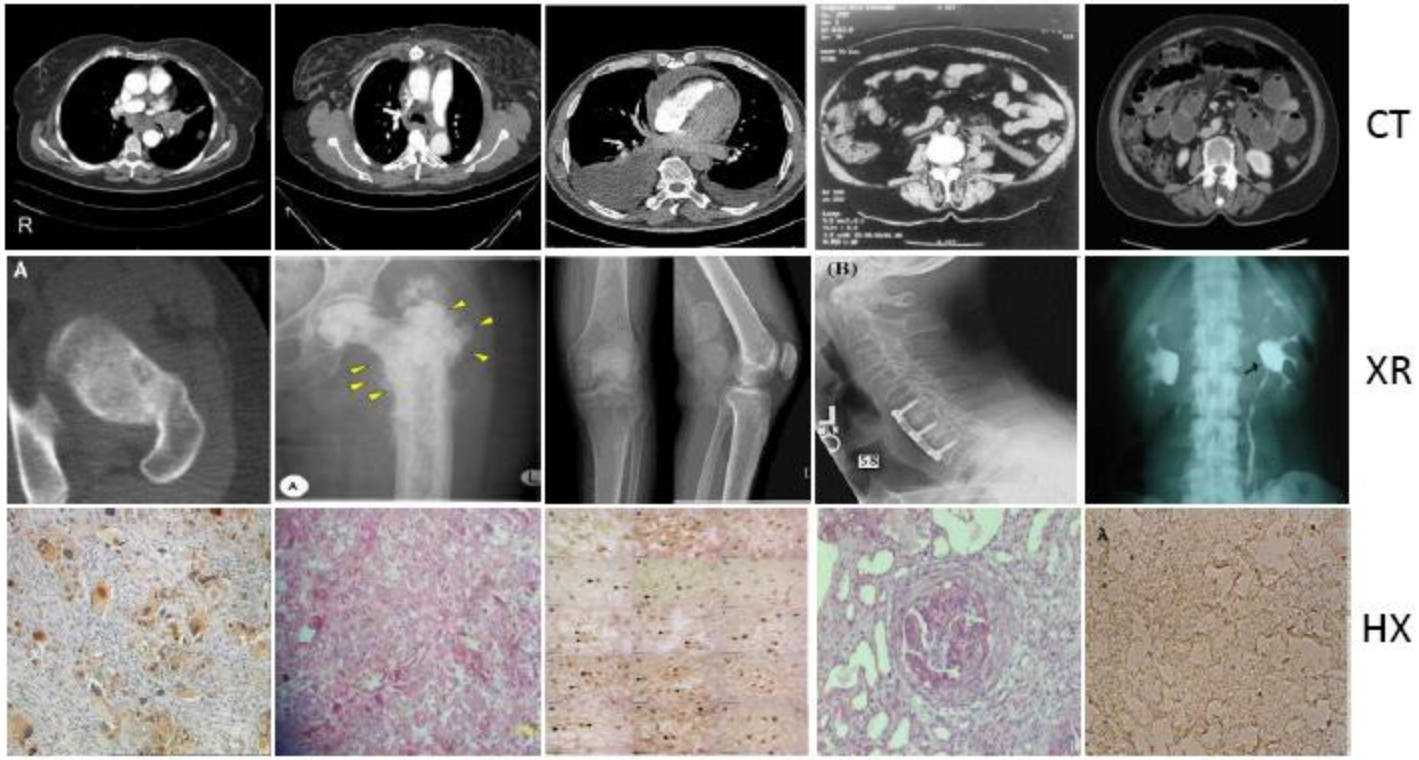


- Constr

	Original	1st	2nd	3rd	4th
VOC	3.53	3.67	3.71	3.72	3.73
HSV	3.17	3.28	3.33	3.34	3.35

Potential Applications

Medical Image Retrieval



Thanks!

Questions and comments