Graph-Based Discriminative Learning for Location Recognition

Song Cao and Noah Snavely

Department of Computer Science, Cornell University

Introduction

- ▸ Goal: Recognize the an image's location by matching to a database
- ▸ Challenges: matching is time consuming; image retrieval is noisy
- ▸ Previous Approaches: image retrieval based & direct matching
- ▸ Our Approach:
  - ▸ **Use an *image graph* to learn local similarity functions**
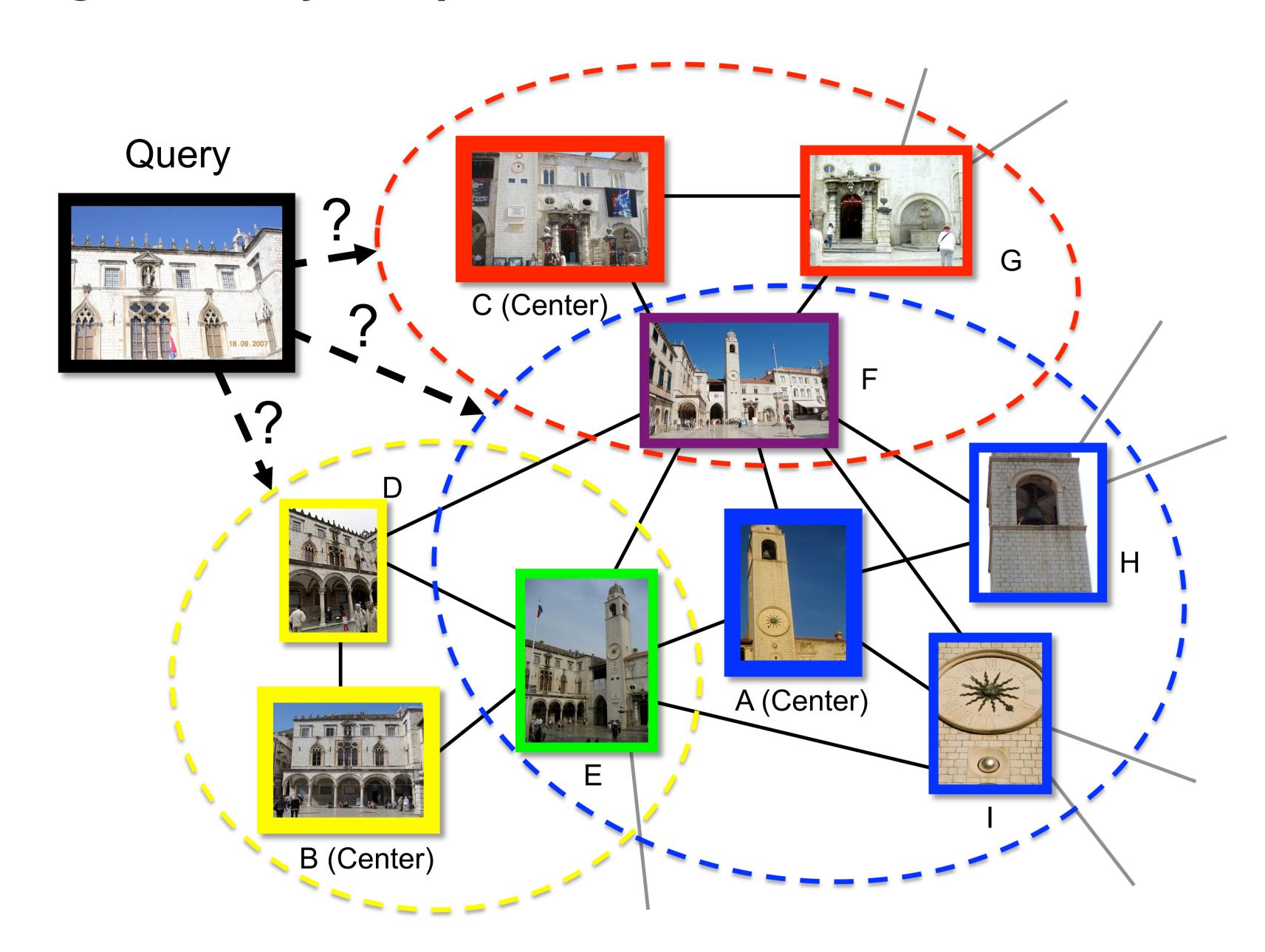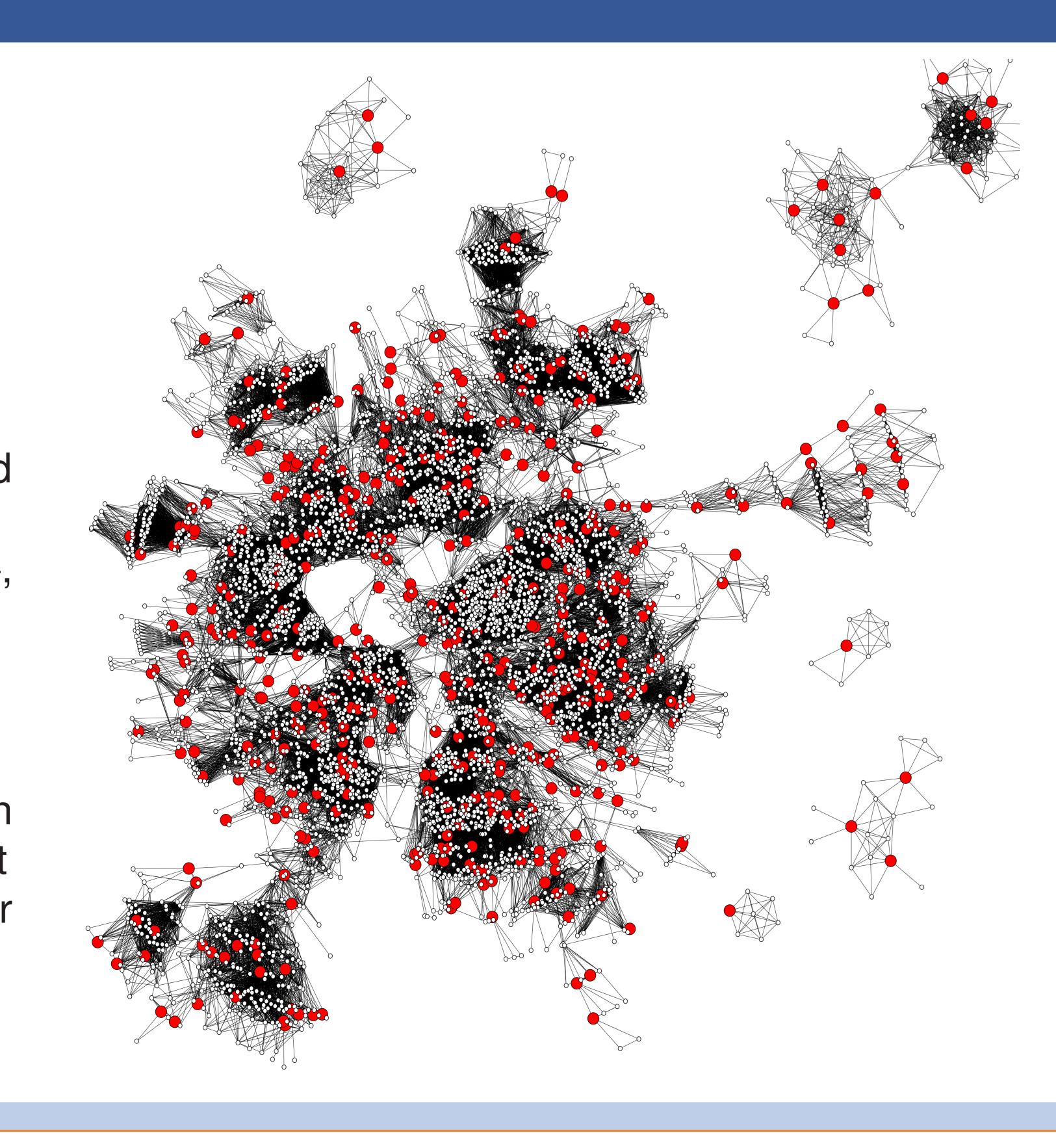 - ▸ **Encourage diversity in top ranked results**

Image Graphs

- ▸ Nodes are images
- ▸ Only geometrically consistent images are connected
- ▸ Edge weights defined by *Jaccard Index* $J(a,b)$: $\frac{N(a,b)}{N(a)+N(b)-N(a,b)}$, and thresholded to improve robustness
- ▸ On the right: an example image graph on Dubrovnik dataset (red nodes are center images selected)

Overview of Approach
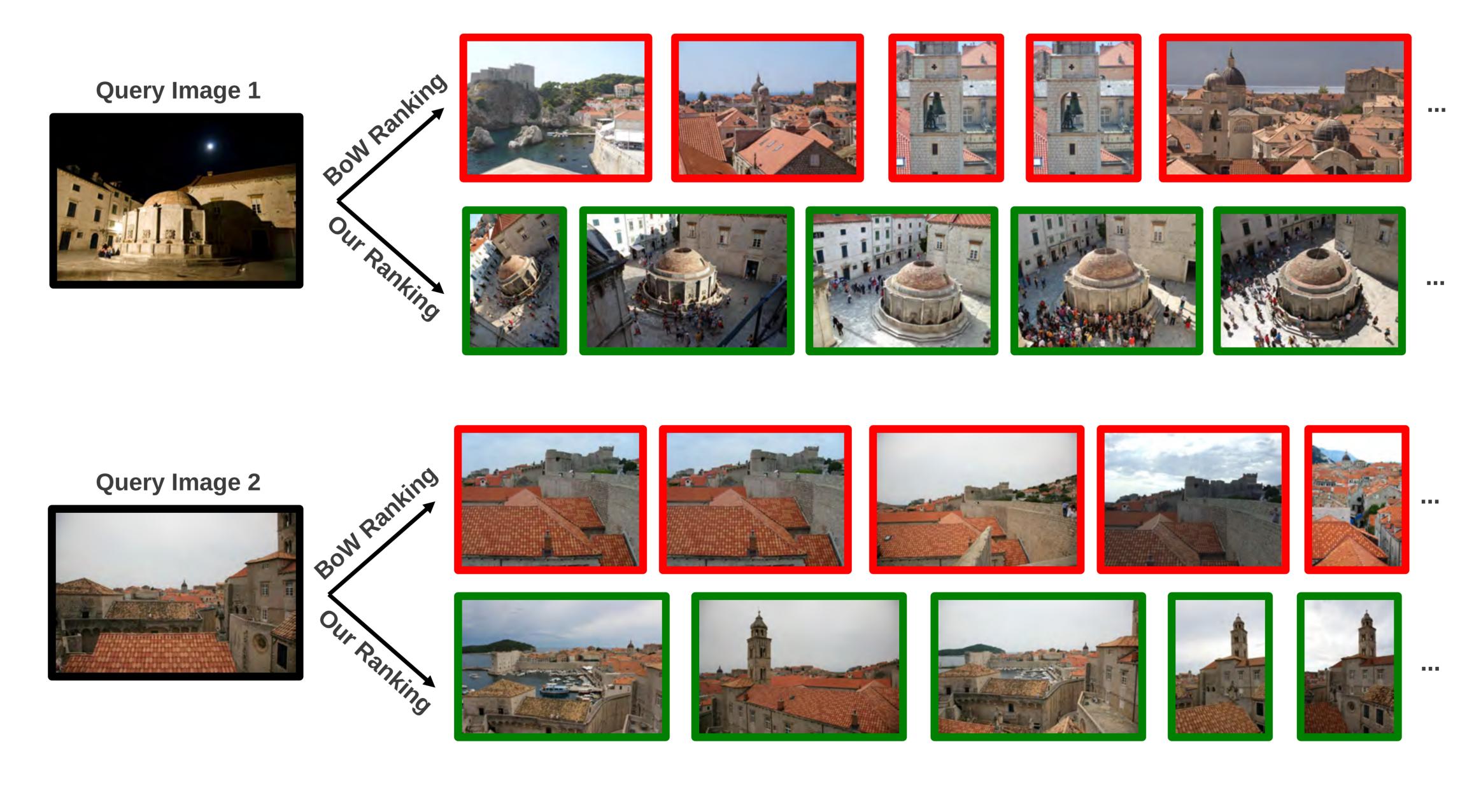
- ▸ Training:
  - ▸ 1. Compute a **covering** of the graph with a set of subgraphs (select center images or neighborhoods in the image graph).
  - ▸ 2. Learn and calibrate an SVM-based distance metric for each subgraph.
- ▸ Testing:
  - ▸ 3. Use the models in step 2 to compute the distance from a query image to each database image, and **generate a ranked shortlist** of possible image matches.
  - ▸ 4. Perform geometric verification sequentially with the top database images in the shortlist.

Generating Ranking Results

- ▸ Ranked neighborhoods are **concatenated** to form a ranking list of all DB images
- ▸ Order within each neighborhood determined by BoW similarity
- ▸ Goal: to have **the first true match** appear in ranked shortlist as early as possible.
- ▸ Comparison of BoW image retrieval ranking and our learned ranking:

- ▸ Ranking can be further improved by enforcing **diversity** in top results: pick the next image **conditioned** on previous one failing to match

Experiments

Dataset	# Queries	# DB Imgs	# Clusters	Ave. Cluster Size
Dubrovnik [1]	800	6,044	188	206.7
Rome [1]	1,000	15,179	352	293.0
Aachen [2]	369	4,479	161	82.0

Table : Top-K accuracies

Dubrovnik (Specific Vocab.)

Method	top1	top2	top5	top10	mAP
BoW	87.50%	92.75%	97.62%	98.50%	0.401
BoW+RR	87.50%	93.38%	96.63%	97.50%	0.058
Co-ocset	87.50%	92.50%	97.50%	98.62%	0.389
GPS Model	87.87%	89.75%	91.75%	93.25%	0.367
Global Model	85.37%	91.63%	95.87%	97.38%	**0.643**
Instance Model	90.00%	95.13%	98.12%	98.50%	0.643
GBP	**94.38%**	96.37%	98.25%	98.50%	0.626
GBP+RR	**94.38%**	96.25%	98.62%	99.13%	0.273
GBP+RR+BoW	94.25%	**97.12%**	**99.37%**	**99.50%**	0.122

Dubrovnik (Generic Vocab.)

Method	top1	top2	top5	top10	mAP
BoW	75.88%	83.00%	90.88%	95.63%	0.512
BoW+RR	75.88%	83.62%	93.25%	**96.25%**	0.065
GBP	81.25%	85.13%	88.13%	90.00%	**0.512**
GBP+RR	81.25%	83.87%	89.88%	95.13%	0.151
GBP+RR+BoW	**81.88%**	**90.00%**	**94.00%**	96.00%	0.085

Rome

Method	top1	top2	top5	top10	mAP
BoW	97.40%	98.50%	99.50%	99.60%	0.674
BoW+RR	97.40%	98.70%	99.10%	99.10%	0.047
GBP	97.80%	98.70%	99.30%	99.30%	**0.789**
GBP+RR	97.80%	**98.80%**	99.30%	99.70%	0.403
GBP+RR+BoW	**97.90%**	**99.00%**	**99.70%**	**99.70%**	0.259

Aachen

Method	top1	top2	top5	top10	mAP
BoW	80.76%	83.47%	86.45%	88.35%	0.431
BoW+RR	80.76%	82.66%	86.45%	88.62%	0.069
GBP	**82.38%**	84.55%	86.72%	88.35%	**0.459**
GBP+RR	**82.38%**	83.74%	87.26%	88.89%	0.205
GBP+RR+BoW	**82.38%**	**84.82%**	**88.08%**	**89.16%**	0.185

Reference

[1] Y. Li, N. Snavely, and D. Huttenlocher. Location recognition using prioritized feature matching. In ECCV, 2010.
[2] T. Sattler, T. Weyand, B. Leibe, and L. Kobbelt. Image retrieval for image-based localization revisited. In BMVC, 2012.