Diffusion for Objects Retrieval

Sung-eui Yoon

https://sgvr.kaist.ac.kr

KAIST

Student Presentation Guidelines

- **Good summary, not full detail, of the paper**
  - **Talk about motivations of the work**
  - **Give a broad background on the related work**
  - **Explain main idea and results of the paper**
  - **Discuss strengths and weaknesses of the method**

- **Prepare an overview slide**
  - **Talk about most important things and connect them well**

High-Level Ideas

- **Deliver most important ideas and results**
  - **Do not talk about minor details**
  - **Give enough background instead**

- **Deeper understanding on a paper is required**
  - **Go over at least two related papers and explain them in a few slides**

- **Spend most time to figure out the most important things and prepare good slides for them**

Deliver Main Ideas of the Paper

- **Identify main ideas/contributions of the paper and deliver them**

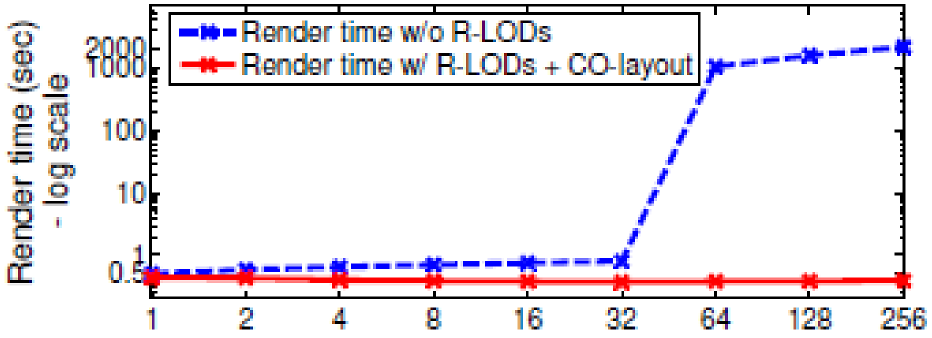
- **If there are prior techniques that you need to understand, study those prior techniques and explain them**
  - **For example, A paper utilizes B's technique in its main idea. In this case, you need to explain B to explain A well.**

Be Honest

- Do not skip important ideas that you don't know
  - Explain as much as you know and mention that you don't understand some parts
- If you get questions you don't know good answers, just say it
- In the end, you need to explain them before the semester ends at KLMS board

Result Presentation

- **Give full experiment settings and present data with the related information**
  - **What does the x-axis mean in the below image?**



- **After showing the data, give a message that we can pull of the data**
- **Show images/videos, if there are**

Utilizing Existing Resources

- **Use author's slides, codes, and video, if they exist**
- **Give proper credits or citations**
  - **Without them, you are cheating!**

Audience feedback form

Date:
Talk title:
Speaker:

1. Was the talk well organized and well prepared?

5: Excellent 4: good 3: okay 2: less than average 1: poor

2. Was the talk comprehensible? How well were important concepts covered?

5: Excellent 4: good 3: okay 2: less than average 1: poor

Any comments to the speaker

Prepare Quiz

- **Review most important concepts of your talk**
  - **Prepare two multiple-choices questions**

- **Example: What is the biased algorithm?**
  - **A: Given N samples, the expected mean of the estimator is I**
  - **B: Given N samples, the exp. Mean of the estimator is I + e**
  - **C: Given N samples, the exp. Mean of the estimator is I + e, where e goes to zero, as N goes to infinite**

- **Grade them in the scale of 0 to 10 and send it to TA**

Class Objectives are:

- Diffusion process
- Hypergraph Propagation and Community Selection for Objects Retrieval

- At last class:
  - Person Re-identification
  - Unsupervised Approaches
  - Part-based Pseudo Label Refinement for Unsupervised Person Re-ID

Challenging in revisited Oxford & Paris

Query

Hard cases

The yellow boxes are the hard cases for the each query.

Hypergraph Propagation and *Community Selection* for *Objects Retrieval*

Guoyuan An, Yuchi Huo, Sung-Eui Yoon

Slide ack: Guoyuan An

Current achievement

Query

Top 20 items of the initial search result; mAP 0.077 (M) and 0.062 (H)

Top 20 items of the result after CS and HD; total 43 SPs; mAP 0.988 (M) and 0.987 (H)

SOTA and the opportunities

- We have the current SOTA result. (NeurIPS 2021)

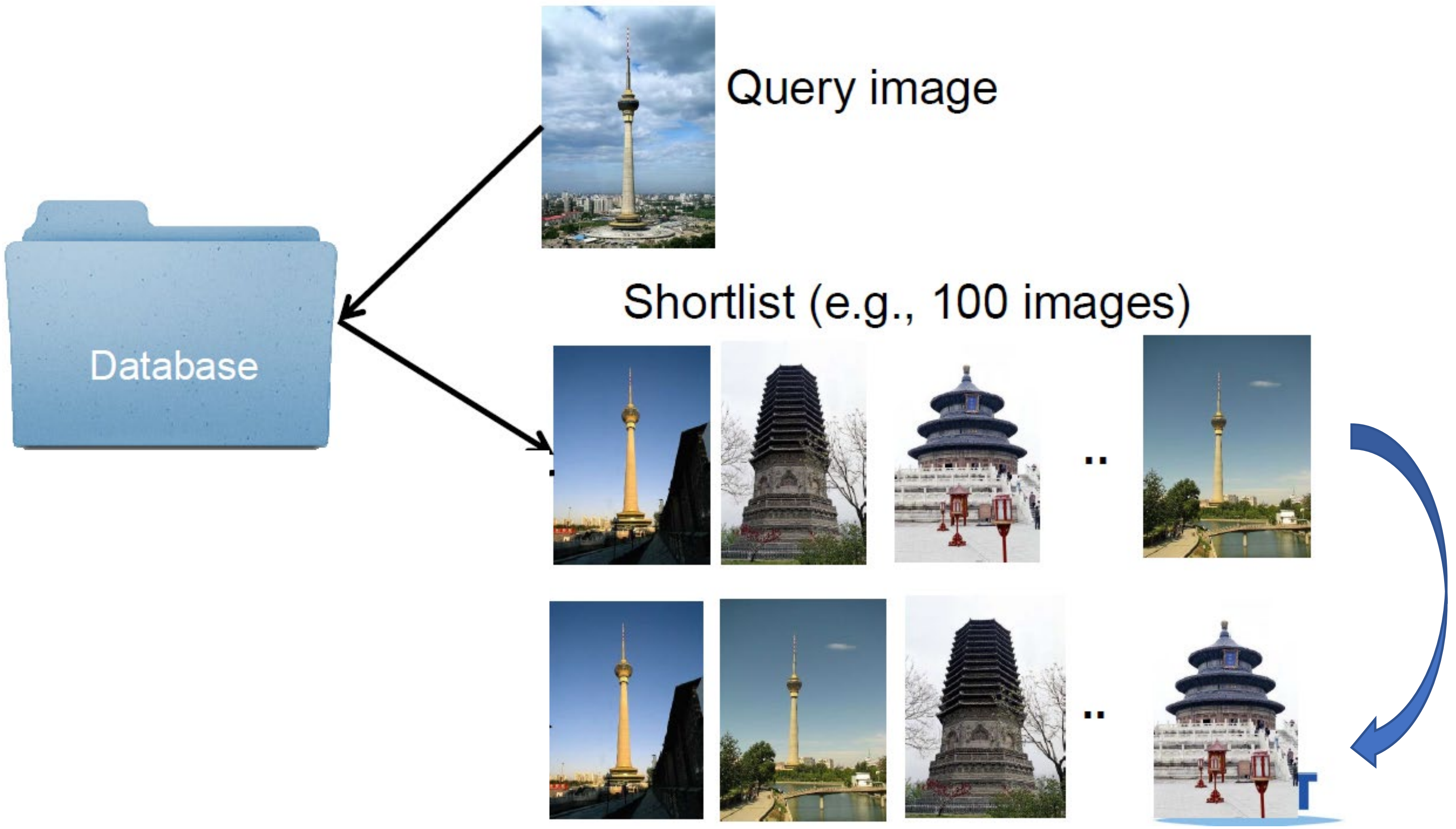
Table 1: Results (% mAP) on the ROxf/RPar datasets and their large-scale versions ROxf+1M/RPar+1M, with both Medium and Hard evaluation protocols.

	ROxf		ROxf+R1M		RPar		RPar+R1M	
Method	M	H	M	H	M	H	M	H
DELG (Global) [5]	76.3	55.6	63.7	37.5	86.6	72.4	70.6	46.9
DELG (Global + Local) [5]	81.2	64.0	69.1	47.5	87.2	72.8	71.5	48.7
Average QE [9]	77.2	57.1	68.5	43.0	87.6	74.3	75.4	54.8
Average QE with decay [13]	78.4	58.0	70.4	44.7	88.2	75.3	76.2	56.0
α QE [27]	65.2	43.2	57.0	30.2	91.0	81.2	81.0	64.1
Diffusion [16]	81.0	59.3	71.5	46.8	91.4	82.7	79.2	64.7
Hypergrah Propagation (Ours)	**85.7**	**70.3**	**78.0**	**60.0**	**92.6**	**83.3**	**86.6**	**72.7**

- There are still a lot challenges / opportunities.
- Check our result to find more opportunities:

Review: pipeline of image search

Pipeline of image retrieval



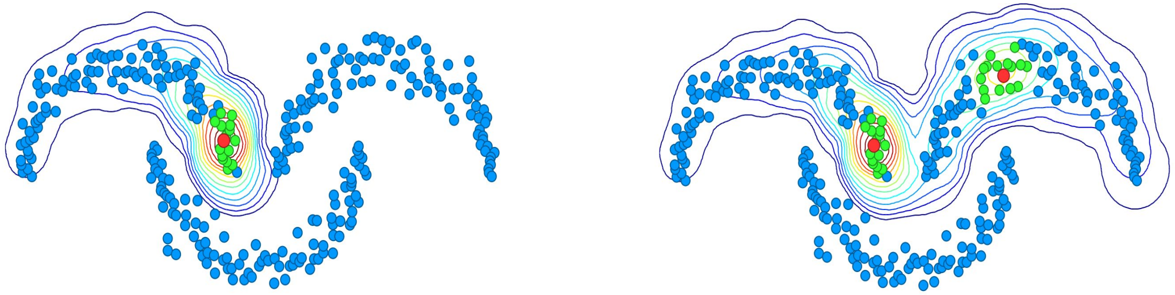
Re-ranking:
1. Diffusion & query expansion
2. Spatial verification

Diffusion and its issue

- We only introduce the high level idea about diffusion here.
- For the detailed explanation, check: https://github.com/anguoyuan/Diffusion-for-retrievl-python.git

Problem

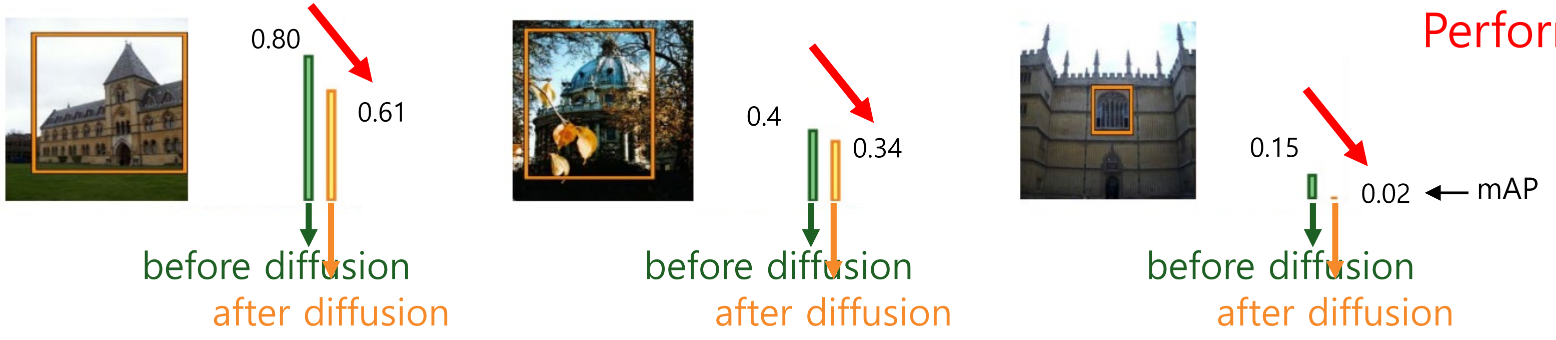
- Diffusion captures the image manifold in the feature space.
- It is a popular and powerful technique to improve the quality of image retrieval.



Iscen, … Efficient Diffusion on Region Manifolds: Recovering Small Objects with Compact CNN Representations, CVPR 2017

- However, some works observed that diffusion degrade the retrieval performance of the hard cases.

Performance degrades



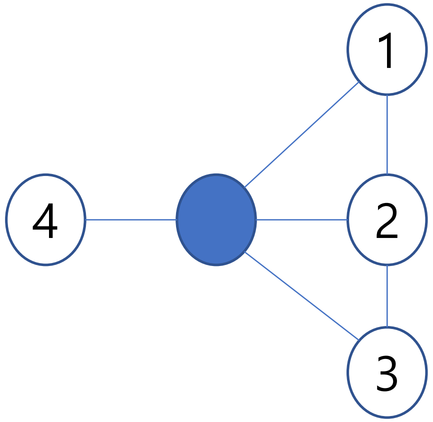
Radenovic, … Efficient Diffusion on Region Manifolds: Recovering Small Objects with Compact CNN Representations, CVPR 2018

How does diffusion work?

- Manifold: even though the images of the sequence contain the same object, the descriptors may be completely unrelated after a certain point.
- Diffusion performs similarity propagation to improve the performance.

How does diffusion work?

- Think about a droplet of ink into the water.
- Node 4 will get larger ranking score (more ink color) than node 1, 2, and 3.
- Same theory with Markov Chain, PageRank,…

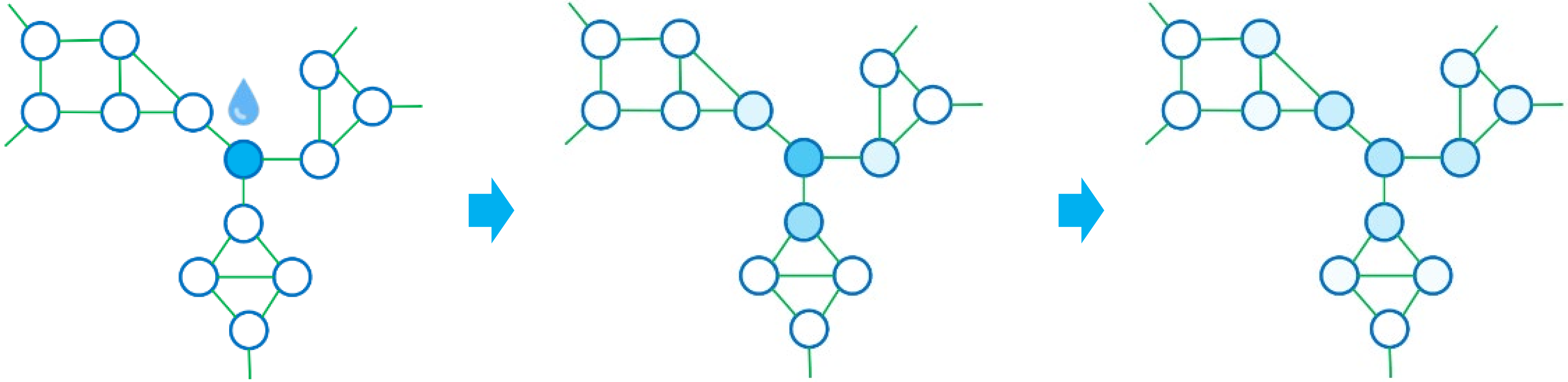


The blue node is the initial query, which is the droplet of ink.

How does diffusion work?

- Diffusion uses the structure information to rank the database images.

- Random walk

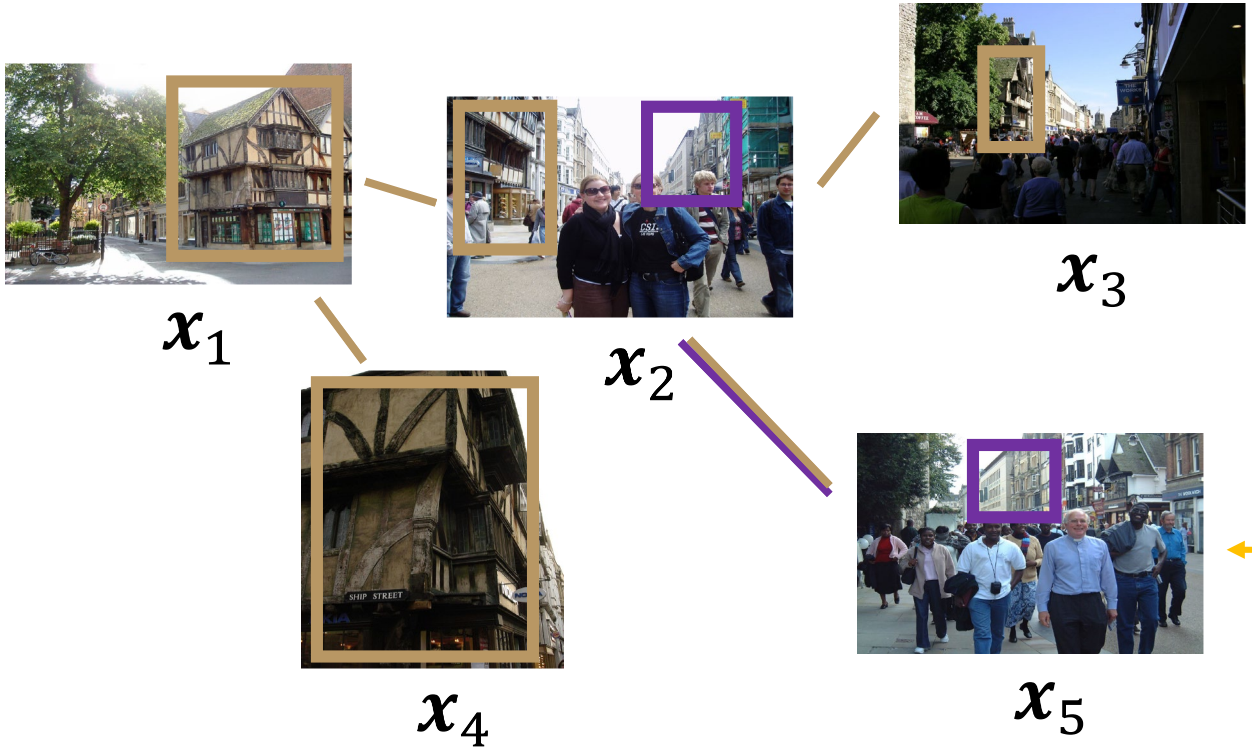


Node: image

Edge: transition probabilities

Where does diffusion fail?

- However, if the dataset is difficult, some images in a manifold may not contain a same object.
- Thus, diffusion inevitably includes false positives.

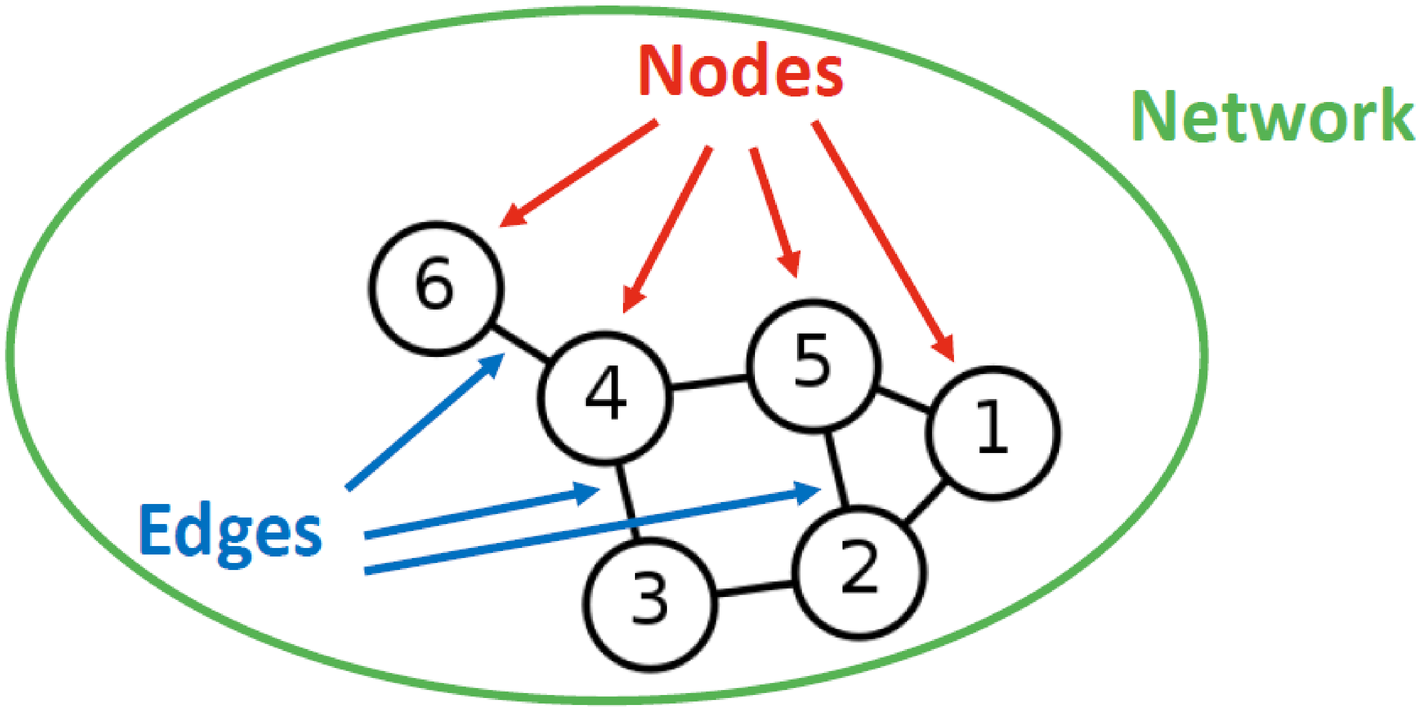


← false positives

Solution: hypergraph propagation

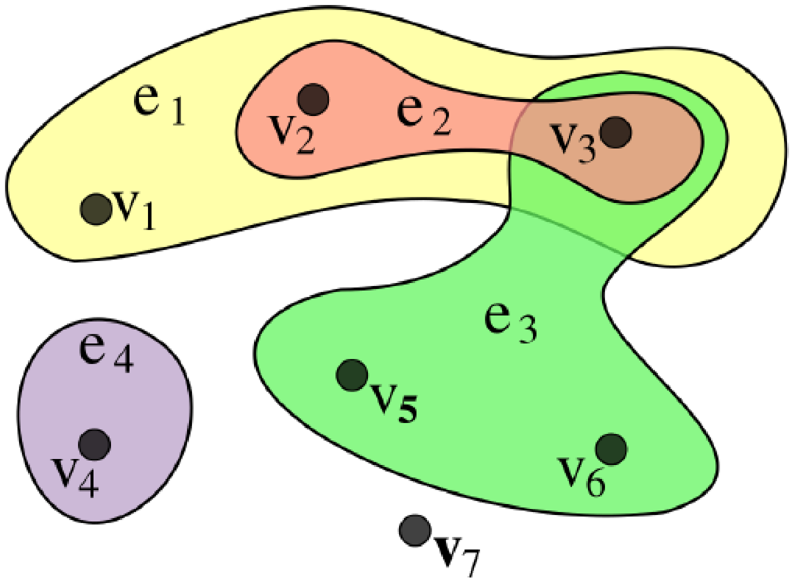
Networks (Graphs)

- A **network** (or a **graph**) G consists of
  - V: set of **nodes (or vertices)**
  - E: set of **edges (or links)**
    - Each edge is associated with a pair of nodes

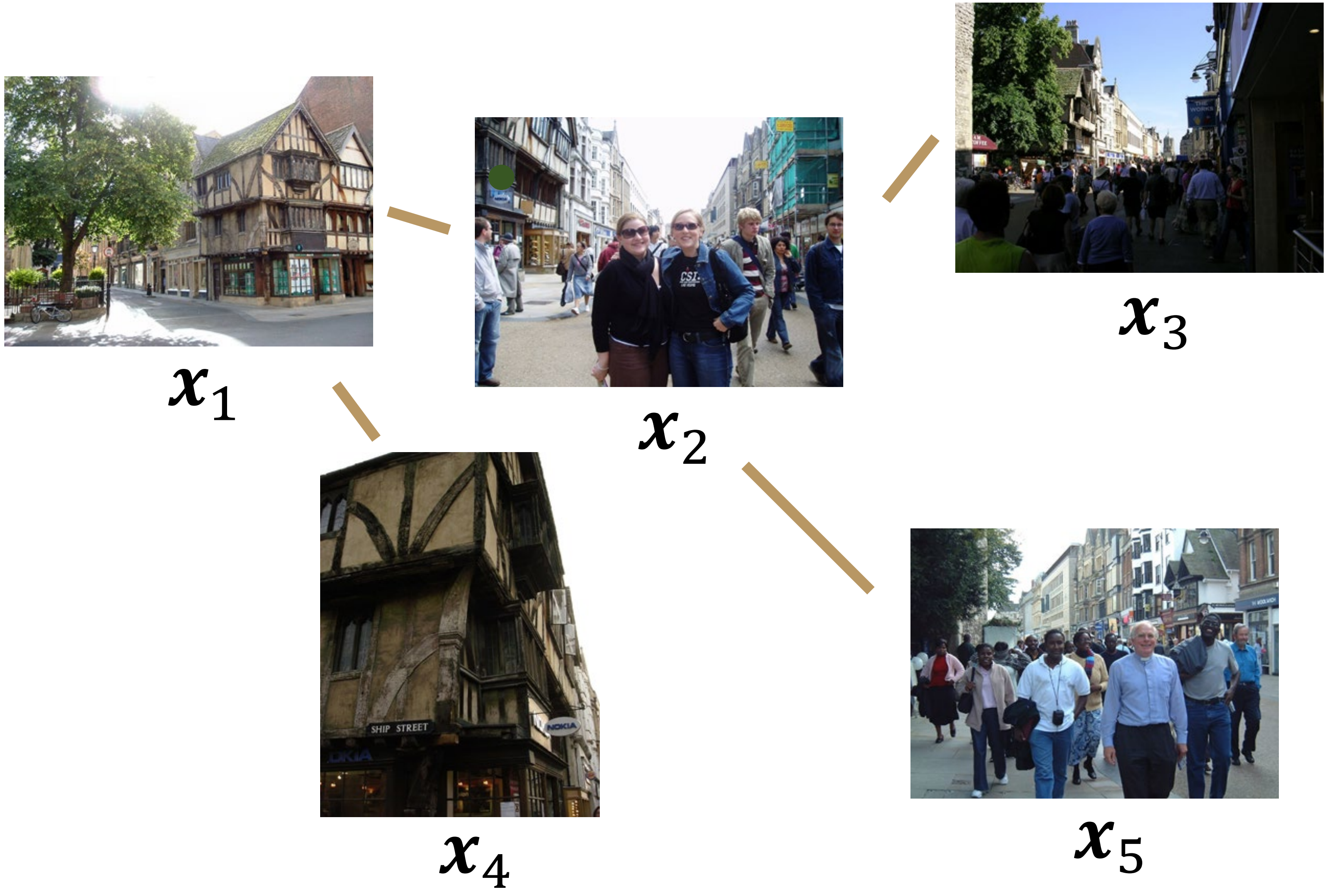


Hypergraphs

- A **hypergraph** (or a **hypernetwork**) G consists of
  - V: set of **nodes (or vertices)**
  - H: set of **hyperedges**
    - Each hyperedge is associated with a non-empty subset of nodes
    - Each hyperedge can contain an **arbitrary** number of nodes

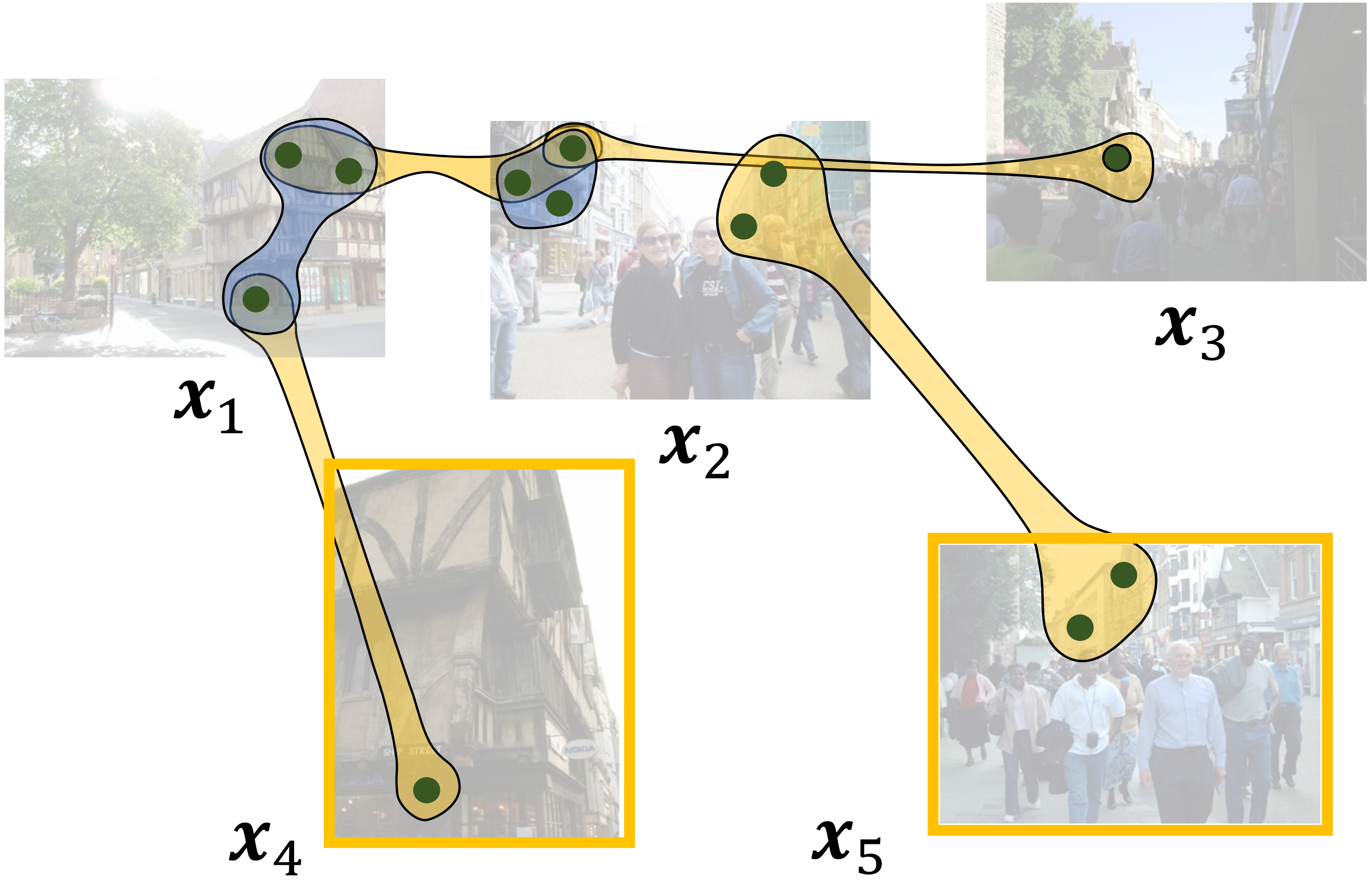


Solution: hypergraph



Solution: hypergraph

- We treat each local feature as a node, and connect these nodes using hyperedges.
- The relation of local features are found by spatial verification.
  - *Recall: two representative reranking methods are diffusion and spatial verification.*



Quantitative result

- Accuracy: significant performance improvement

Table 1: Results (% mAP) on the ROxf/RPar datasets and their large-scale versions ROxf+1M/RPar+1M, with both Medium and Hard evaluation protocols.

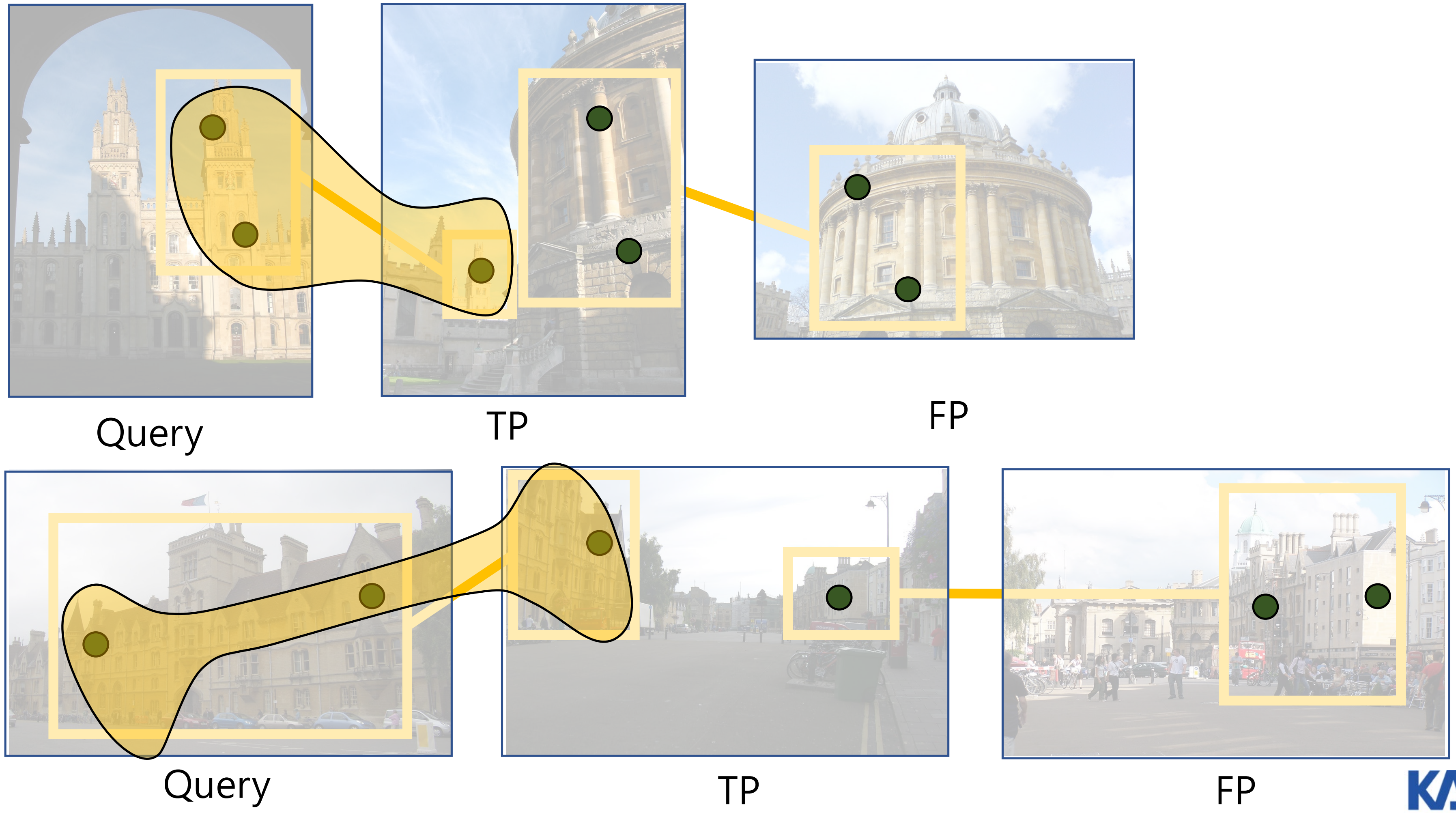
	ROxf		ROxf+R1M		RPar		RPar+R1M	
Method	M	H	M	H	M	H	M	H
DELG (Global) [5]	76.3	55.6	63.7	37.5	86.6	72.4	70.6	46.9
DELG (Global + Local) [5]	81.2	64.0	69.1	47.5	87.2	72.8	71.5	48.7
Average QE [9]	77.2	57.1	68.5	43.0	87.6	74.3	75.4	54.8
Average QE with decay [13]	78.4	58.0	70.4	44.7	88.2	75.3	76.2	56.0
α QE [27]	65.2	43.2	57.0	30.2	91.0	81.2	81.0	64.1
Diffusion [16]	81.0	59.3	71.5	46.8	91.4	82.7	79.2	64.7
Hypergrah Propagation (Ours)	**85.7**	**70.3**	**78.0**	**60.0**	**92.6**	**83.3**	**86.6**	**72.7**

- Speed: significant performance improvement

initial search	hypergraph propagation	uncertainty calculation	spatial verification
0.62 s	1.07 s	0.0003 s	41.12 s

Qualitative result

- For each query, we correctly only accept the true positive (TP) and reject the false positive (FP) in the initial ranking list.



Problem of spatial verification

Estimate the accuracy of initial search

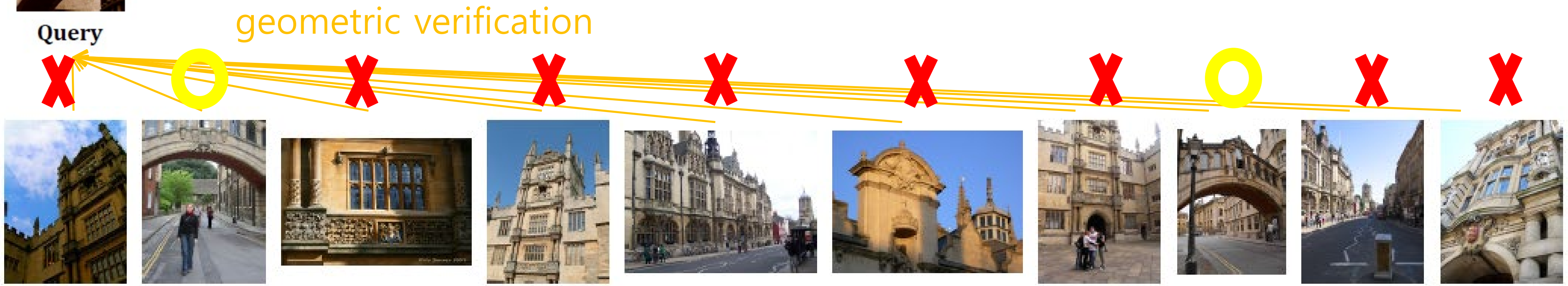
- Estimating the accuracy of initial search is important for
  - Reranking with diffusion
  - Search engine improvement
  - User experience.
- Traditional way is doing the geometric verification for the top 100 searched images.

(b) Breakdown of average time per query.

very slow

initial search	hypergraph propagation	uncertainty calculation	spatial verification
0.62 s	1.07 s	0.0003 s	$\boxed{41.12\ \text{s}}$

Query



Solution: community selection

Predict the initial search quality

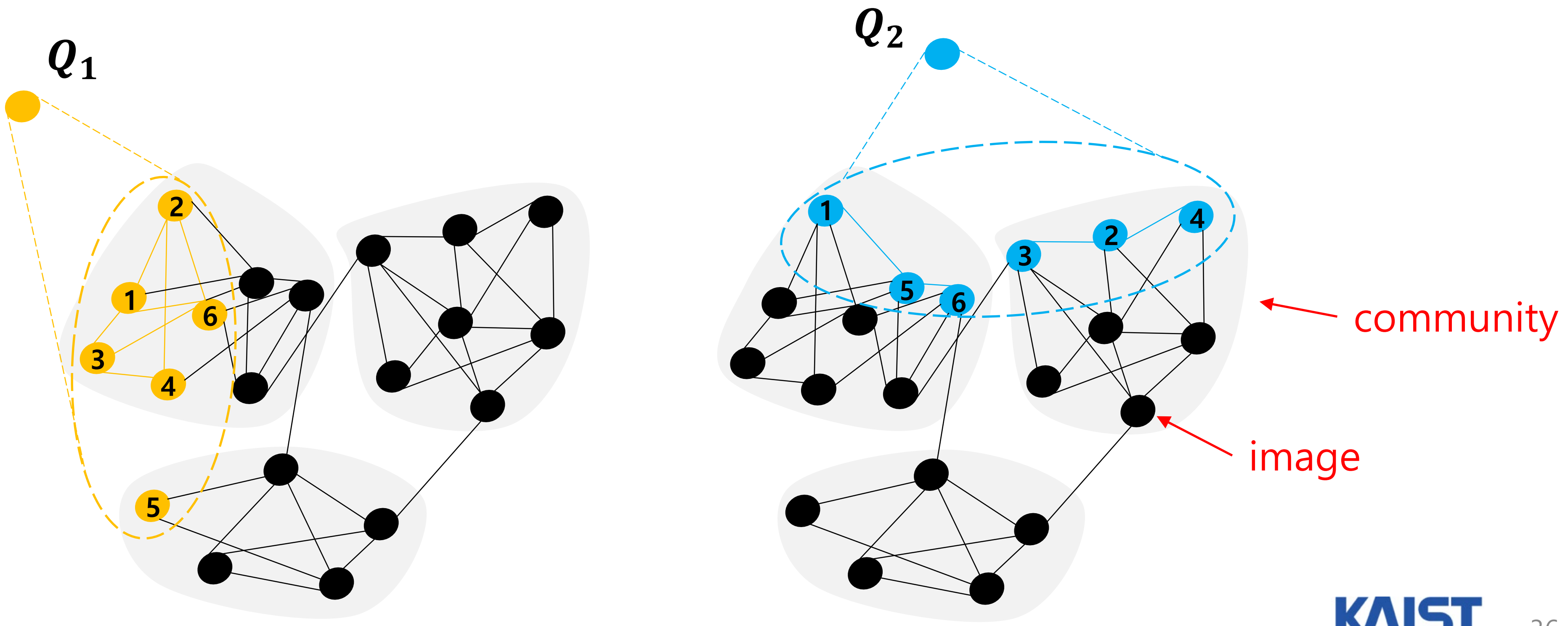
- Reranking, especially spatial verification, is time consuming.
- However, not all initial rankings need reranking. If we can predict the initial search quality in advance, we can only do the reranking for low-quality initial search and directly accept the high-quality initial search.
- We propose community selection technique to predict the search quality.
  - Not that this process requires no *ground truth* or *user feedback*.
  - This process is very fast.

(b) Breakdown of average time per query.

initial search	hypergraph propagation	uncertainty calculation	spatial verification
0.62 s	1.07 s	0.0003 s	41.12 s

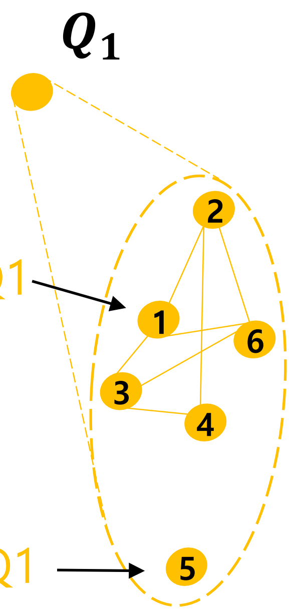
Graph approach: community selection

- We observe that the images containing a same object usually belong to a same community.
- By checking whether the searched images are in the same community, we can evaluate the quality of the initial search without query time spatial verification.



How to find the community?

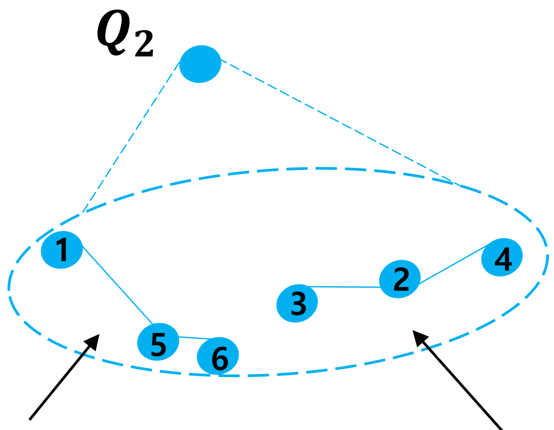
- It is not necessary detect all communities in the database in advance.
- Instead, we detect the connected components among the top K (K=6 in the follow figure) of each query; number in node shows searched ranking
- We then calculate the entropy w/ numbers of nodes of the component.
  - Higher entropy means higher uncertainty and low quality.



component 1 for Q1

Component 2 for Q1

Entropy is low



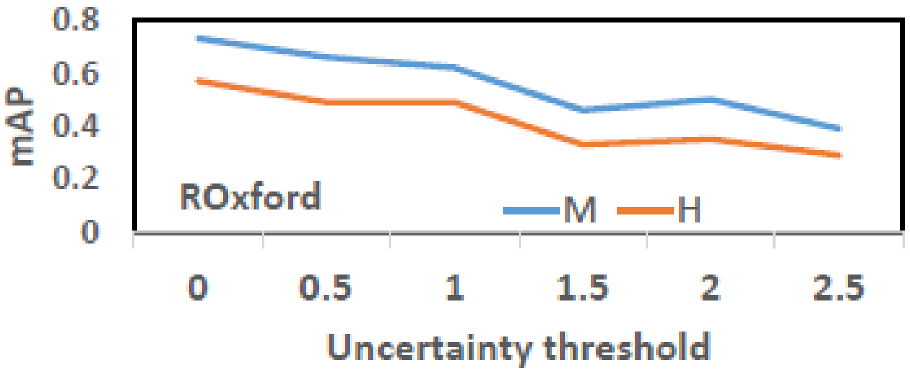
component 1 for Q1

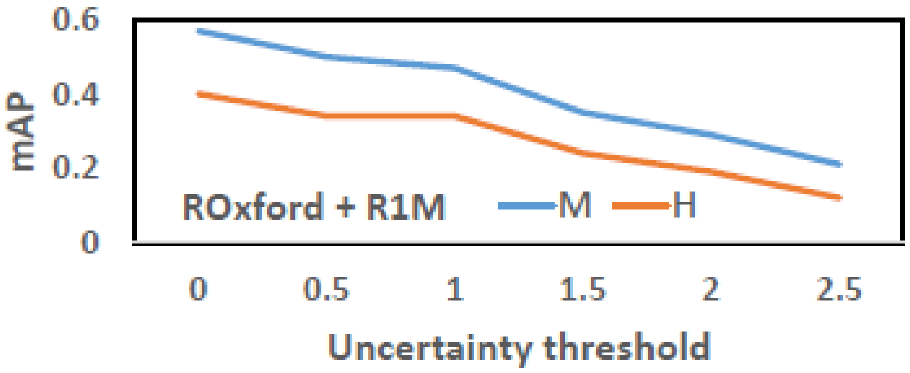
Component 2 for Q1

Entropy is high

Quantitative result

- The uncertainty index predicts the quality of initial search without spatial verification or any user feedback.





- The speed of calculating the uncertainty is very fast.

(b) Breakdown of average time per query.

initial search	hypergraph propagation	uncertainty calculation	spatial verification
0.62 s	1.07 s	0.0003 s	41.12 s

Thank you

Detail and code: https://sgvr.kaist.ac.kr/~guoyuan/hypergraph_propagation/

Class Objectives were:

- **Diffusion process**
- **Hypergraph Propagation and Community Selection for Objects Retrieval**

Summary

- **Locally invariant features**
  - **Key point localization and Harris detector**

- **Scale invariant region selection**
  - **Automatic scale selection**
  - **Laplacian of Gradients (LoG) ≈ Difference of Gradients (DoG)**
  - **SIFT as a local descriptor**

Summary

- CNN based image descriptors
  - Training losses, data, and benchmarks
- Re-ranking and fast k-NN search
  - Spatial verification and query expansion
  - Inverted index and inverted multi-index
- Hashing techniques using hyperplanes and hyperspheres
  - Sematic hashing using deep learning
- Person Re-identification
  - Unsupervised approaches and part-based pseudo label refinement
- Diffusion process
  - Hypergraph propagation and community selection

KAIST