ICML
International Conference
On Machine Learning

# Cluster-Aware Similarity Diffusion for Instance Retrieval


Jifei Luo[1], Hantao Yao[2] and Changsheng Xu[2]

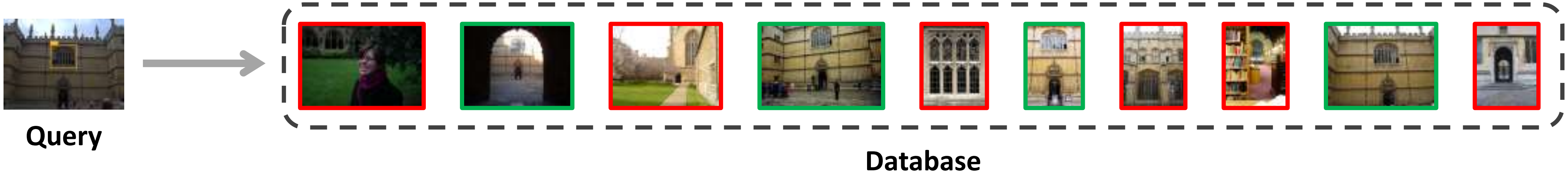
[1]University of Science and Technology of China
[2]Institute of Automation, Chinese Academy of Sciences

# Background

- Instance retrieval task aims to search through a large scale database to find the relevant images that share similar content.

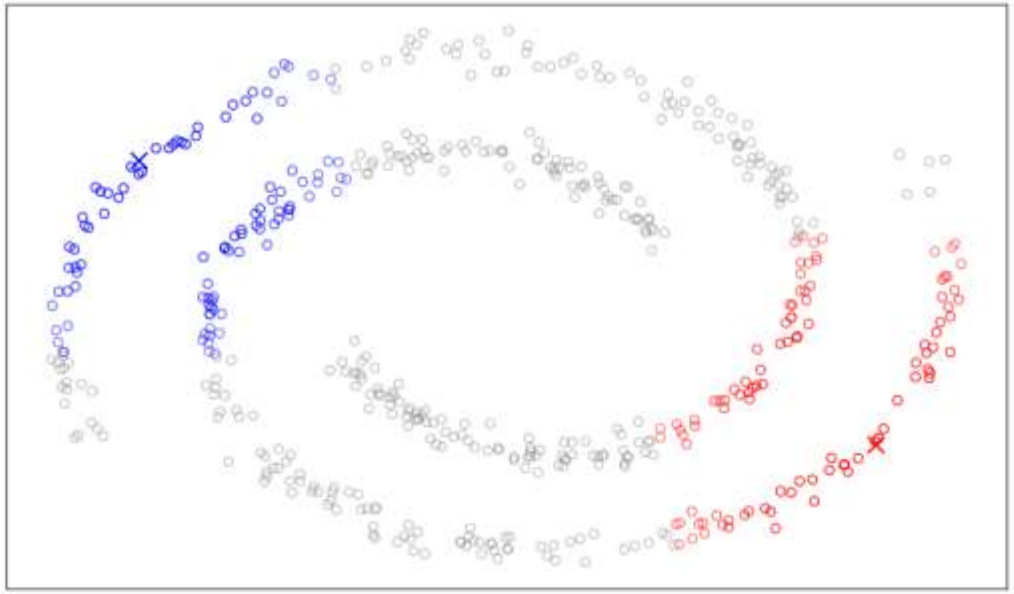
# Background

- Instance retrieval task aims to search through a large scale image database to find the relevant images that share similar content.
- However, the global descriptors may not yield out the optimal retrieval performance.

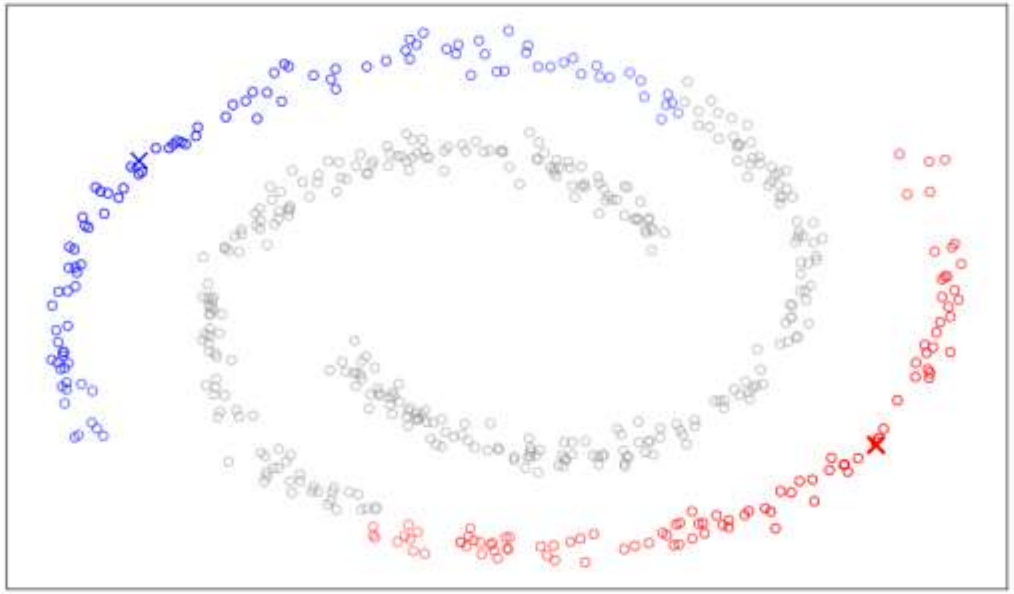


# Motivation

- Here is a toy example on Swiss Roll dataset.
- The retrieval results returned by Euclidean distance and the idea retrieval results based on manifold structure.



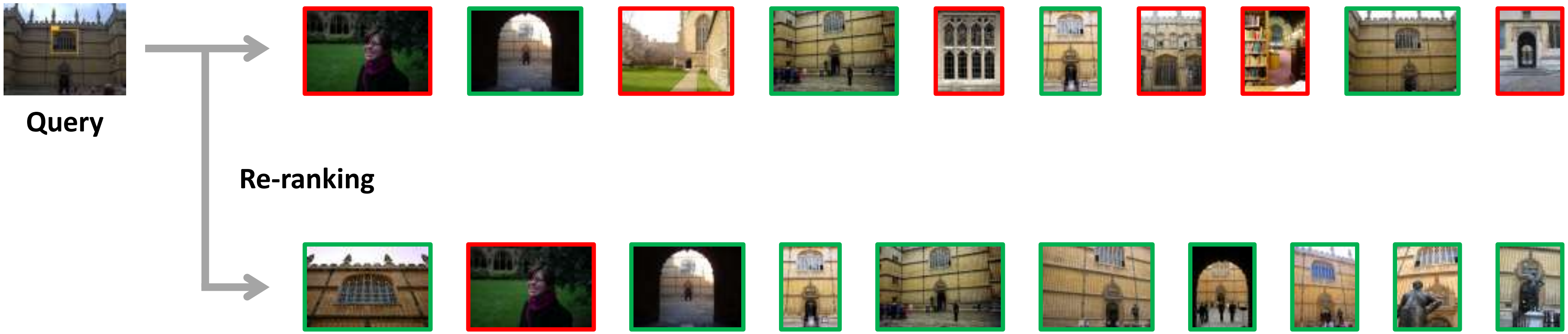
Euclidean distance



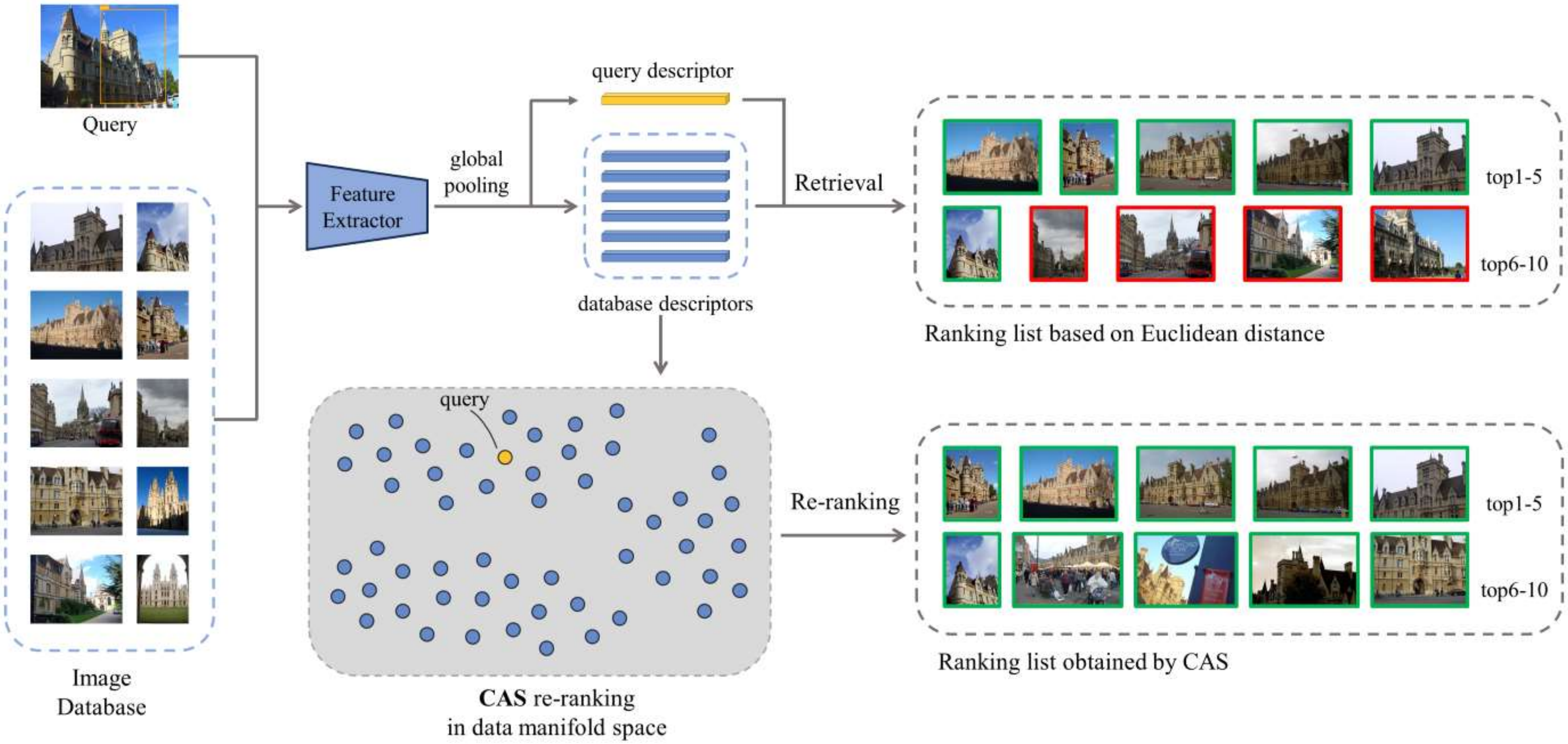
Ideal results

# Motivation

- The features of similar images are more likely to be lying on the same manifold.
- Based on this assumption, we can leverage the underlying manifold information to re-rank the initial retrieval list to achieve better performance.



# Workflow


Query
Image
Database
Feature
Extractor
global
pooling
query descriptor
database descriptors
Retrieval
top1-5
top6-10
Ranking list based on Euclidean distance
query
CAS re-ranking
in data manifold space
Re-ranking
top1-5
top6-10
Ranking list obtained by CAS

# Methodology

- We first construct an affinity graph to model the underlying data manifold structure,

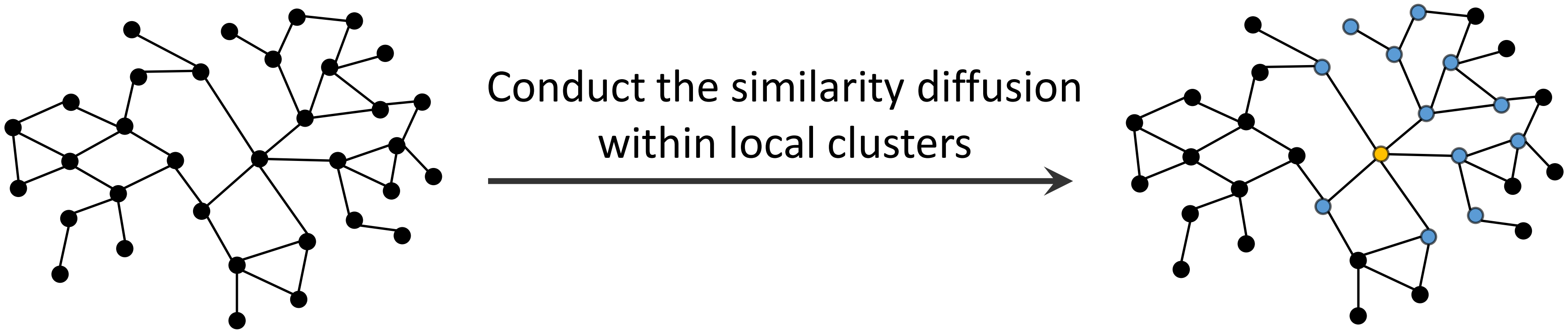
$$\boldsymbol{W}_{ij} = \mathbf{1}_{ij} \exp\left(-d^2(i,j)/\sigma^2\right)$$

where $\mathbf{1}$ is an indicator that represents which vertices are connected.

# Methodology

- We introduce a novel Bidirectional Similarity Diffusion strategy and constrain the diffusion in the local cluster $\mathcal{C}$ explicitly.

$$\min_{\boldsymbol{F}} \frac{1}{4}\sum_{k=1}^{n}\sum_{i,j=1}^{n} \boldsymbol{W}_{ij}\left(\frac{\boldsymbol{F}_{ki}}{\sqrt{\boldsymbol{D}_{ii}}} - \frac{\boldsymbol{F}_{kj}}{\sqrt{\boldsymbol{D}_{jj}}}\right)^2 + \boldsymbol{W}_{ij}\left(\frac{\boldsymbol{F}_{ik}}{\sqrt{\boldsymbol{D}_{ii}}} - \frac{\boldsymbol{F}_{jk}}{\sqrt{\boldsymbol{D}_{jj}}}\right)^2 + \mu\|\boldsymbol{F} - \boldsymbol{E}|_F^2$$



# Methodology

- We introduce a novel Bidirectional Similarity Diffusion strategy and constrain the diffusion in the local cluster $\mathcal{C}$ explicitly.

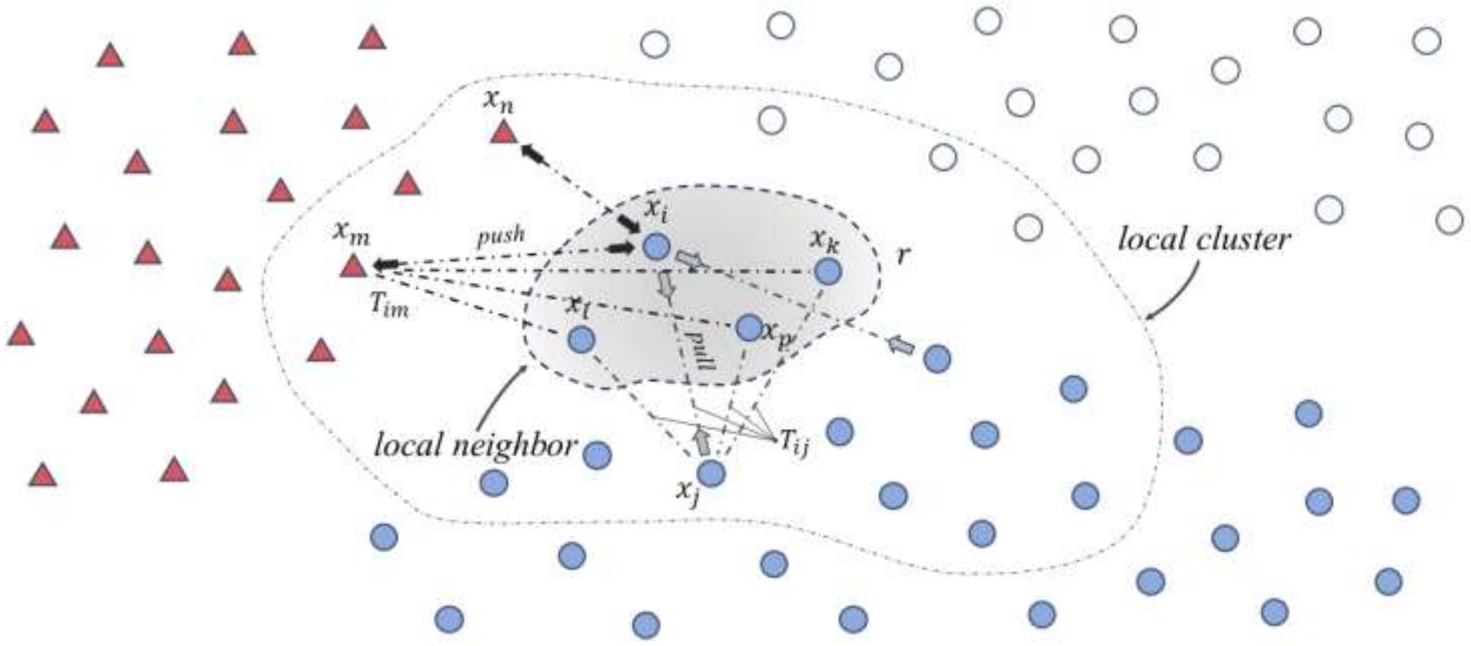
$$\min_{\boldsymbol{F}} \frac{1}{4}\sum_{k=1}^{n}\sum_{i,j=1}^{n}\boldsymbol{W}_{ij}\left(\frac{\boldsymbol{F}_{ki}}{\sqrt{\boldsymbol{D}_{ii}}}-\frac{\boldsymbol{F}_{kj}}{\sqrt{\boldsymbol{D}_{jj}}}\right)^2+\boldsymbol{W}_{ij}\left(\frac{\boldsymbol{F}_{ik}}{\sqrt{\boldsymbol{D}_{ii}}}-\frac{\boldsymbol{F}_{jk}}{\sqrt{\boldsymbol{D}_{jj}}}\right)^2+\mu\|\boldsymbol{F}-\boldsymbol{E}|_F^2$$

- The optimal result is equivalent to the solution of the following Lyapunov equation, which can be iteratively solved.

$$(\boldsymbol{I}-\alpha\boldsymbol{S})\boldsymbol{F}+\boldsymbol{F}(\boldsymbol{I}-\alpha\boldsymbol{S})=2(1-\alpha)\boldsymbol{E}$$

# Method

- To ensure the similarity consistency from instance to local neighbors, Neighbor-guided Similarity Smooth is further introduced to refine matrix $\boldsymbol{F}$



# Method

- To ensure the similarity consistency from instance to local neighbors, Neighbor-guided Similarity Smooth is further introduced to refine matrix $\boldsymbol{F}$

- Moreover, the similarity within the local neighborhood can be aggregated to form a neighbor-enhanced representation $\widetilde{\boldsymbol{F}}$, and an extra propagation can be applied

$$\widetilde{\boldsymbol{F}}_i = \Big(\kappa \sum_{j \in \xi[i]} \hat{\boldsymbol{F}}_j / |\xi| + \sum_{j \in \mathcal{N}(i,k_2)} \hat{\boldsymbol{F}}_j / k_2\Big) / (\kappa + 1)$$

$$\boldsymbol{F}'_i = \sum_j \boldsymbol{P}_{ij} \widetilde{\boldsymbol{F}}_j$$

# Experiments

- We conduct experiments on both *R*Oxford and *R*Paris datasets.

| Method | Medium | | | | Hard | | | |
|---|---|---|---|---|---|---|---|---|
| | *R*Oxf | *R*Oxf+1M | *R*Par | *R*Par+1M | *R*Oxf | *R*Oxf+1M | *R*Par | *R*Par+1M |
| R-GeM | 67.3 | 49.5 | 80.6 | 57.4 | 44.2 | 25.7 | 61.5 | 29.8 |
| AQE | 72.3 | 56.7 | 82.7 | 61.7 | 48.9 | 30.0 | 65.0 | 35.9 |
| αQE | 69.7 | 53.1 | 86.5 | 65.3 | 44.8 | 26.5 | 71.0 | 40.2 |
| DQE | 70.3 | 56.7 | 85.9 | 66.9 | 45.9 | 30.8 | 69.9 | 43.2 |
| AQEwD | 72.2 | 56.6 | 83.2 | 62.5 | 48.8 | 29.8 | 65.8 | 36.6 |
| LAttQE | 73.4 | 58.3 | 86.3 | 67.3 | 49.6 | 31.0 | 70.6 | 42.4 |
| kNN | 71.3 | 54.7 | 83.8 | 63.2 | 49.1 | 29.2 | 66.4 | 36.7 |
| DFS | 72.9 | 59.4 | 89.7 | 74.0 | 50.1 | 34.9 | 80.4 | 56.9 |
| FSR | 72.7 | 59.6 | 89.6 | 73.9 | 49.6 | 34.8 | 80.2 | 56.7 |
| RDP | 75.2 | 55.0 | 89.7 | 70.0 | 58.8 | 33.9 | 77.9 | 48.0 |
| EIR | 74.9 | **61.6** | 89.7 | 73.7 | 52.1 | 36.9 | 79.8 | 56.1 |
| GSS | 78.0 | 61.5 | 88.9 | 71.8 | 60.9 | 38.4 | 76.5 | 50.1 |
| EGT | 74.7 | 60.1 | 87.9 | 72.6 | 51.1 | 36.2 | 76.6 | 51.3 |
| SSR | 74.2 | 54.6 | 82.5 | 60.0 | 53.2 | 29.3 | 65.6 | 35.0 |
| CSA | 78.2 | 61.5 | 88.2 | 71.6 | 59.1 | 38.2 | 75.3 | 51.0 |
| SG | 71.4 | 53.9 | 83.6 | 61.5 | 49.5 | 28.8 | 67.6 | 35.8 |
| **CAS** | **80.7** | **61.6** | **91.0** | **75.5** | **64.8** | **39.1** | **80.7** | **59.7** |

# Experiments

- We conduct experiments on both *R*Oxford and *R*Paris datasets.

| Method | Easy | | Medium | | Hard | |
|---|---|---|---|---|---|---|
| | *R*Oxf | *R*Par | *R*Oxf | *R*Par | *R*Oxf | *R*Par |
| MAC | 47.2 | 69.7 | 34.6 | 55.7 | 14.3 | 32.6 |
| AQE | 54.4 | 80.9 | 40.6 | 67.0 | 17.1 | 45.2 |
| αQE | 50.3 | 77.8 | 37.1 | 64.4 | 16.3 | 43.0 |
| DQE | 50.1 | 78.1 | 37.8 | 66.5 | 16.0 | 45.7 |
| kNN | 56.6 | 79.7 | 41.6 | 66.5 | 17.4 | 44.5 |
| AQEwD | 52.8 | 79.6 | 39.7 | 65.0 | 17.3 | 42.9 |
| DFS | 54.6 | 83.8 | 40.6 | 74.0 | 18.8 | 58.1 |
| FSR | 54.4 | 83.9 | 40.4 | 73.5 | 18.4 | 57.5 |
| EIR | 57.9 | 86.9 | 44.2 | 76.8 | 22.2 | 60.5 |
| RDP | 59.0 | 85.2 | 45.3 | 76.3 | 21.4 | 58.9 |
| GSS | 60.0 | 87.5 | 45.4 | 76.7 | 22.8 | 59.7 |
| **CAS** | **68.6** | **90.1** | **52.9** | **82.3** | **30.4** | **68.1** |

| Method | Easy | | Medium | | Hard | |
|---|---|---|---|---|---|---|
| | *R*Oxf | *R*Par | *R*Oxf | *R*Par | *R*Oxf | *R*Par |
| R-MAC | 61.2 | 79.3 | 40.2 | 63.8 | 10.1 | 38.2 |
| AQE | 69.4 | 85.7 | 47.8 | 71.1 | 15.9 | 47.9 |
| αQE | 64.9 | 84.7 | 42.8 | 70.8 | 11.4 | 47.8 |
| DQE | 65.5 | 84.9 | 45.3 | 71.9 | 15.5 | 49.1 |
| kNN | 70.6 | 84.6 | 48.9 | 70.2 | 16.0 | 46.1 |
| AQEwD | 70.5 | 85.9 | 48.7 | 70.7 | 15.3 | 46.9 |
| DFS | 70.0 | 87.5 | 51.8 | 78.8 | 20.3 | 63.5 |
| FSR | 69.7 | 87.3 | 51.4 | 78.1 | 20.1 | 62.6 |
| EIR | 68.0 | 89.4 | 50.8 | 78.7 | 21.7 | 63.3 |
| RDP | 73.7 | 88.8 | 54.3 | 79.6 | 22.2 | 61.3 |
| GSS | 75.0 | 89.9 | 54.7 | 78.5 | 24.4 | 60.5 |
| **CAS** | **82.6** | **90.0** | **62.5** | **82.5** | **34.1** | **67.4** |

# Experiments

- Here we exhibit some quantitative reranking results on *R*Oxford dataset.

# Time Complexity

- The overall time complexity of CAS is $\mathcal{O}(n^3)$.
- For tasks that require finding neighbors in a manifold space, our method can serve as a plug-in module to enhance overall performance.

| Method | Time Complexity | Re-ranking Latency (ms) |
|---|---|---|
| αQE | $\mathcal{O}(n^2 d)$ | 121 |
| $k$-recip | $\mathcal{O}(n^3)$ | 8,524 |
| DFS | $\mathcal{O}(n^3)$ | 1,857 |
| RDP | $\mathcal{O}(n^3)$ | 6,018 |
| GSS | $\mathcal{O}(n^3)$ | >5 min |
| CAS | $\mathcal{O}(n^3)$ | 1,278 |

# Conclusion

- Our proposed CAS effectively mitigate the negative influence from nearby manifold and achieves superior performance.
- In our future work, we plan to utilize the concept of optimal transportation and apply the inference stage within the affinity graph to fully exploit the underlying manifold information.

# Cluster-Aware Similarity Diffusion for Instance Retrieval

Jifei Luo[1], Hantao Yao[2] and Changsheng Xu[2]

[1]University of Science and Technology of China
[2]Institute of Automation, Chinese Academy of Sciences