细粒度场景级基于草图的图像检索方法

SceneSketcher: Fine-Grained Image Retrieval with Scene Sketches

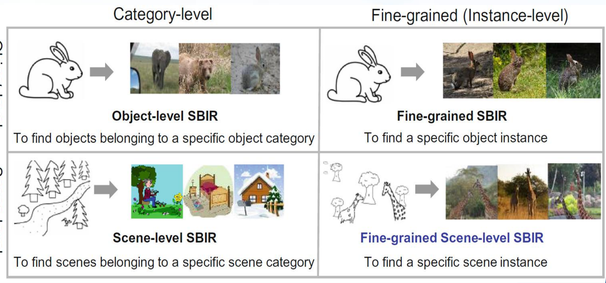
Fang Liu Changqing Zou Xiaoming Deng* Ran Zuo
Yu-Kun Lai Cuixia Ma* Yong-Jin Liu* Hongan Wang

European Conference on Computer Vision (**ECCV**) 2020

cuixia@iscas.ac.cn

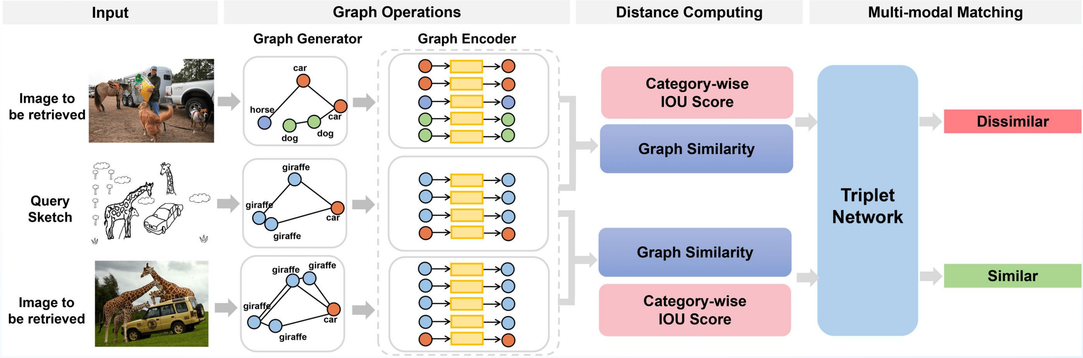
Abstract

- In this paper, for the first time, we study the fine-grained scene-level SBIR problem which aims at retrieving scene images satisfying the user's specific requirements via a freehand scene sketch.
- We propose a graph embedding based method to learn the similarity measurement between images and scene sketches, which models the multi-modal information, including the size and appearance of objects as well as their layout information.



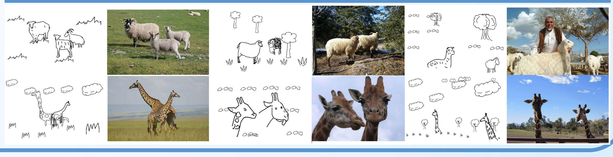
Methods

Our proposed framework consists of **three phases**:



① **graph operations;**
② **distance computing;**
③ **multi-modal matching.**

Examples of scene-level sketch dataset.

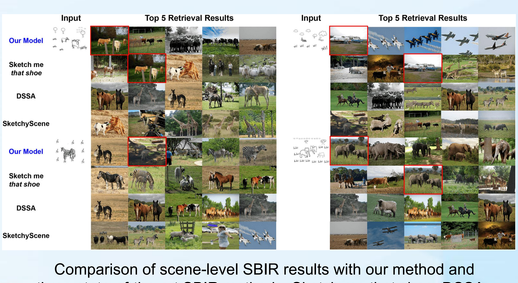


Results - Comparison with SoTA

Comparison with Baselines on our database (210 testing images) and our extended database (5210 testing images):

- Our model achieves significantly higher recall than the other baselines.

	Our sketch database			Extended database		
	Recall@1	Recall@5	Recall@10	Recall@10	Recall@50	Recall@100
HOG+BoW+RankSVM [20]	0.48	1.43	4.76	0.48	0.48	0.48
Dense HOG+RankSVM [36]	0.48	3.81	5.71	0	0.95	1.91
Sketch-a-Net+RankSVM [37]	0.48	3.33	4.76	0	0.95	2.86
Sketch me that shoe [36]	6.19	17.15	32.86	1.90	6.19	8.57
DSSA [27]	0.48	3.81	7.62	0	0.95	1.90
SketchyScene [40]	1.43	4.76	8.57	0.48	0.95	2.86
Our model	**31.91**	**66.67**	**86.19**	**38.10**	**68.10**	**82.86**

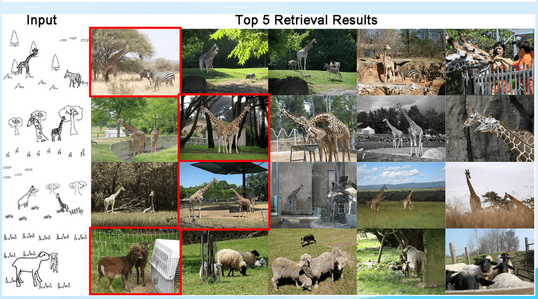


Comparison of scene-level SBIR results with our method and three state-of-the-art SBIR methods: Sketch me that shoe, DSSA , SketchyScene. The ground truth matches are highlighted with red rectangles.

Results - Ablation Study

- Visual features, category labels and position information all contribute to enhancing the retrieval performance.
- Category-wise IoU is important for our network

Model settings	Our sketch database			Extended database		
	Recall@1	Recall@5	Recall@10	Recall@1	Recall@5	Recall@10
1. Visual feature as graph only	24.29	51.90	77.14	8.09	18.09	25.23
2. Category label as graph only	29.52	62.86	82.38	8.57	20.95	30.00
3. Visual feature and category label as graph	30.48	64.76	83.81	11.43	23.33	30.95
4. Graph triplet loss only	13.33	30.00	47.62	2.38	6.67	10.00
5. Category-wise IoU only	28.10	61.90	80.0	6.67	19.05	24.29
6. $IoU_{category}$ only	23.82	59.05	76.19	4.76	16.19	23.81
7. Global IoU only	5.24	19.05	28.10	0	0.48	2.38
8. $IoU_{category}$+Graph feature	24.76	59.05	78.57	4.76	16.67	23.81
9. Our model	**31.91**	**66.67**	**86.19**	**12.38**	**26.67**	**38.10**



Examples of scene sketch SBIR results on our extended database (5210 test images).