Global Features are All You Need for Image Retrieval and Reranking
Supplementary Material

Shihao Shao[1 *†] Kaifeng Chen[2] Arjun Karpur[2] Qinghua Cui[1] André Araujo[2] Bingyi Cao[2 *†]

[1]Peking University [2]Google Research

S1. Reranking more candidates in SuperGlobal

This section serves as an extension to Section 5.3 and further evaluates using SuperGlobal with additional candidates on Revisited Oxford 5k (+1M) and Revisited Paris 6k (+1M) [3]. As presented in Table S1, SuperGlobal achieves further performance improvements when reranking additional images. It significantly outperforms CVNet reranking by 6.9% on Revisited Oxford+1M Hard, and surpasses SuperGlobal (rerank top 400) by 3.2% in the same dataset. Even with additional candidates, SuperGlobal (rerank top 1600) enjoys significant latency gains over CVNet reranking (rerank top 400).

S2. Parameter study for each module

Our method consists of several modules that are sensitive to the choice of parameter. Therefore, one important aspect of our work is to seek the optimal values for the core parameters of each component. In this section, we verify the validity of these values by conducting grid searches on different parameter values. When appropriate, we present two digits after the decimal due to the minor differences in values.

p for GeM+. As mentioned in Section 5.1, we used $\mathcal{R}$Oxford 5k [3] to estimate p and obtain the value of 4.6 for GeM+. We show the results of different p values in Table S2 to verify that our p is optimal.

p_{ms} for Scale-GeM. Here we perform grid search to explore the influence of p_{ms} in Scale-GeM. The results are detailed in Table S3.

p_r for Regional-GeM. Regional GeM consists of L_p pool and GeM+. Table S4 shows how the p_r value of L_p pool affects retrieval performance.

ReLU threshold. Here we present the study of the relationship between the threshold α of ReLU and retrieval performance. We conduct grid search to investigate the optimal α and summarize the results in Table S5.

S3. Combining SuperGlobal with other state-of-the-art models.

SuperGlobal can easily be adopted to existing retrieval methods for further improvements. Table S6 demonstrates that adopting SuperGlobal modules (GeM+, Scale-GeM, and Regional-GeM) and further performing SuperGlobal reranking on the DELG [1] pretrained weights outperforms CVNet reranking [2].

S4. Generalizing SuperGlobal reranking

SuperGlobal proposes the idea to rerank by further improving global feature of images via feature aggregation. This idea can be generalized when combined with other global features, *e.g.*, DELG-Global [1], DOLG [4] or CVNet-Global [2]. Here, we evaluate retrieval performance when applying SuperGlobal reranking on top of CVNet-Global. Please note that the other modules introduced in SuperGlobal (e.g. GeM+, Scale-GeM, Regional-GeM) are not included in this section of experiments. As shown in Table S7, we report that applying SuperGlobal reranking module to CVNet-Global significantly improve the performance in both $\mathcal{R}$Oxford and $\mathcal{R}$Paris datasets. When comparing with CVNet reranking (Table S1), using SuperGlobal reranking still shows superior performance in the $\mathcal{R}$Paris dataset.

References

[1] Bingyi Cao, Andre Araujo, and Jack Sim. Unifying deep local and global features for image search. In *ECCV*, 2020. 1, 2, 3

[2] S. Lee, H. Seong, S. Lee, and E. Kim. Correlation Verification for Image Retrieval. In *Proc. CVPR*, 2022. 1, 2, 3

[3] Filip Radenović, Ahmet Iscen, Giorgos Tolias, Yannis Avrithis, and Ondřej Chum. Revisiting oxford and paris: Large-scale image retrieval benchmarking. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018. 1, 2, 3

[4] Min Yang, Dongliang He, Miao Fan, Baorong Shi, Xuetong Xue, Fu Li, Errui Ding, and Jizhou Huang. Dolg: Single-stage image retrieval with deep orthogonal fusion of local and global features. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021. 1, 2

Method	Medium				Hard			
	$\mathcal{R}$Oxf	$\mathcal{R}$Oxf+1M	$\mathcal{R}$Par	$\mathcal{R}$Par+1M	$\mathcal{R}$Oxf	$\mathcal{R}$Oxf+1M	$\mathcal{R}$Par	$\mathcal{R}$Par+1M
Global feature retrieval								
RN50-DELG [1]	73.6	60.6	85.7	68.6	51.0	32.7	71.5	44.4
RN101-DELG [1]	76.3	63.7	86.6	70.6	55.6	37.5	72.4	46.9
RN50-DOLG [4]	80.5	**76.6**	89.8	80.8	58.8	52.2	77.7	62.8
RN101-DOLG [4]	81.5	77.4	91.0	83.3	61.1	54.8	80.3	66.7
RN50-CVNet [2]	81.0	72.6	88.8	79.0	62.1	50.2	76.5	60.2
RN101-CVNet [2]	80.2	74.0	90.3	80.6	63.1	53.7	79.1	62.2
RN50-SuperGlobal **[ours]**	**83.9**	74.7	**90.5**	**81.3**	**67.7**	**53.6**	**80.3**	**65.2**
RN101-SuperGlobal **[ours]**	**85.3**	**78.8**	**92.1**	**83.9**	**72.1**	**61.9**	**83.5**	**69.1**
Global feature retrieval + Local feature reranking								
RN50-DELG (GV rerank top 100) [1]	78.3	67.2	85.7	69.6	57.9	43.6	71.0	45.7
RN101-DELG (GV rerank top 100) [1]	81.2	69.1	87.2	71.5	64.0	47.5	72.8	48.7
RN50-CVNet (Rerank top 400) [2]	87.9	80.7	90.5	82.4	75.6	65.1	80.2	67.3
RN101-CVNet (Rerank top 400) [2]	87.2	81.9	91.2	83.8	75.9	67.4	81.1	69.3
SuperGlobal feature retrieval and reranking								
RN50-SuperGlobal (Rerank top 400) **[ours]**	88.8	80.0	92.0	83.4	77.1	64.2	84.4	68.7
RN101-SuperGlobal (Rerank top 400) **[ours]**	90.9	84.4	93.3	84.9	80.2	71.1	86.7	71.4
RN50-SuperGlobal (Rerank top 800) **[ours]**	**88.9**	81.3	93.0	85.4	**77.4**	67.0	86.2	**75.4**
RN101-SuperGlobal (Rerank top 800) **[ours]**	**91.2**	85.5	94.1	86.5	**80.7**	73.5	88.2	74.6
RN50-SuperGlobal (Rerank top 1600) **[ours]**	88.9	**82.0**	**93.3**	**86.8**	76.9	**68.2**	**86.4**	75.0
RN101-SuperGlobal (Rerank top 1600) **[ours]**	91.2	**85.9**	**94.2**	**87.7**	80.6	**74.3**	**88.4**	**77.0**

Table S1: **Comparison to the state-of-the-art methods in image retrieval tasks.** Results (% mAP) on the $\mathcal{R}$Oxford and $\mathcal{R}$Paris datasets[3] (and their large-scale versions $\mathcal{R}$Oxf+1M and $\mathcal{R}$Par+1M), with both Medium and Hard evaluation protocols. Our SuperGlobal retrieval framework outperforms state-of-the-art image retrieval methods by a large margin for every measure. The best scores for RN50 and RN101, with and without reranking, are highlighted in **bold black** and **bold blue**, respectively.

Method	p	Medium		Hard	
		$\mathcal{R}$Oxf	$\mathcal{R}$Par	$\mathcal{R}$Oxf	$\mathcal{R}$Par
SuperGlobal	4.2	90.9	93.4	80.1	86.7
	4.4	90.9	93.4	80.1	86.8
	4.6	90.9	93.3	80.2	86.7
	4.8	91.0	92.3	80.3	86.7
	5.0	90.8	93.3	80.0	86.7

Table S2: Results (% mAP) of conducting grid search on different GeM+ p values on the $\mathcal{R}$Oxford and $\mathcal{R}$Paris datasets [3], with both Medium and Hard evaluation protocols.

Method	p_{ms}	Medium		Hard	
		$\mathcal{R}$Oxf	$\mathcal{R}$Par	$\mathcal{R}$Oxf	$\mathcal{R}$Par
SuperGlobal	1.0	89.5	93.1	77.8	86.2
	1.5	89.5	93.1	77.9	86.2
	2.0	89.7	93.1	78.1	86.2
	2.5	90.4	93.1	79.1	86.2
	3.0	90.6	93.1	79.3	86.3
	$+\infty$	90.9	93.3	80.2	86.7

Table S3: Results (% mAP) of conduct grid search on different Scale-GeM p_{ms} values on the $\mathcal{R}$Oxford and $\mathcal{R}$Paris datasets [3], with both Medium and Hard evaluation protocols.

Method	p_r	Medium		Hard	
		$\mathcal{R}$Oxf	$\mathcal{R}$Par	$\mathcal{R}$Oxf	$\mathcal{R}$Par
SuperGlobal	2.0	90.9	93.3	80.2	86.7
	2.2	90.9	93.3	80.2	86.7
	2.4	90.9	93.3	80.2	86.7
	2.6	90.8	93.3	80.0	86.7
	2.8	90.8	93.4	80.0	86.7

Table S4: Results (% mAP) of conducting grid search on different Regional-GeM p_r values on the $\mathcal{R}$Oxford and $\mathcal{R}$Paris datasets [3], with both Medium and Hard evaluation protocols.

Method	α	Medium		Hard	
		$\mathcal{R}$Oxf	$\mathcal{R}$Par	$\mathcal{R}$Oxf	$\mathcal{R}$Par
SuperGlobal	0.012	90.7	93.3	79.8	86.7
	0.014	90.9	93.3	80.2	86.7
	0.016	90.9	93.4	80.0	86.7
	0.018	90.9	93.4	80.2	86.7
	0.020	90.8	93.3	80.1	86.6

Table S5: Results (% mAP) of conducting grid search on different ReLU threshold on the $\mathcal{R}$Oxford and $\mathcal{R}$Paris datasets [3], with both Medium and Hard evaluation protocols.

Method	Medium		Hard	
	$\mathcal{R}$Oxf	$\mathcal{R}$Par	$\mathcal{R}$Oxf	$\mathcal{R}$Par
RN101-DELG[1]	76.3	86.6	55.6	72.4
RN101-DELG+SuperGlobal pooling [one-stage]	80.0	90.6	60.0	79.8
RN101-DELG+SuperGlobal pooling and reranking (top 400)	88.4	93.1	77.3	86.8

Table S6: Results (% mAP) of adopting SuperGlobal to make further improvement on DELG [1] on the $\mathcal{R}$Oxford and $\mathcal{R}$Paris datasets [3], with both Medium and Hard evaluation protocols.

Method	SuperGlobal (Rerank top 400)	Medium		Hard	
		$\mathcal{R}$Oxf	$\mathcal{R}$Par	$\mathcal{R}$Oxf	$\mathcal{R}$Par
RN101-CVNet-Global[2]	✘	80.2	90.3	63.1	79.1
	✓	83.7	91.6	68.6	82.5

Table S7: Results (% mAP) of adopting SuperGlobal (only reranking) on CVNet-Global [2] on the $\mathcal{R}$Oxford and $\mathcal{R}$Paris datasets [3], with both Medium and Hard evaluation protocols.