Adversarial Ranking Attack and Defense

Mo Zhou, Zhenxing Niu, Le Wang, Qilin Zhang, Gang Hua

May 2020

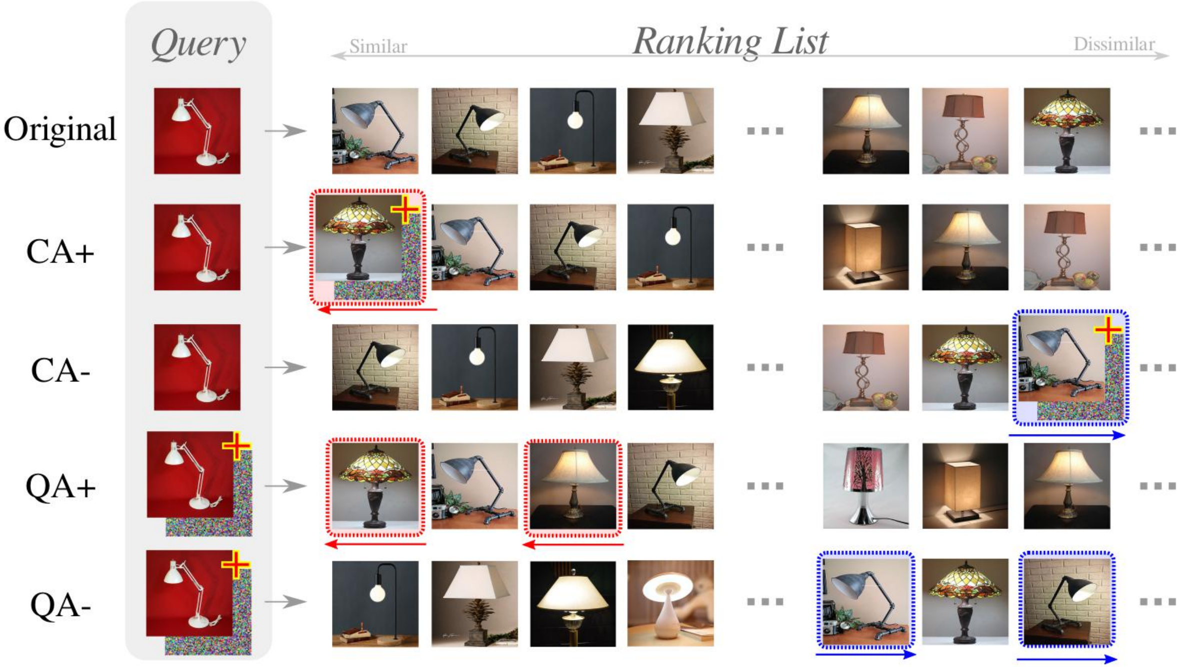
https://arxiv.org/abs/2002.11293

Adversarial Ranking Attack

- **Definition:** Raise or lower the rank of chosen candidates with respect to a specific query set

- Candidate Attack (CA): Raise (CA+) or lower (CA-) the rank by perturbing <u>candidates.</u>

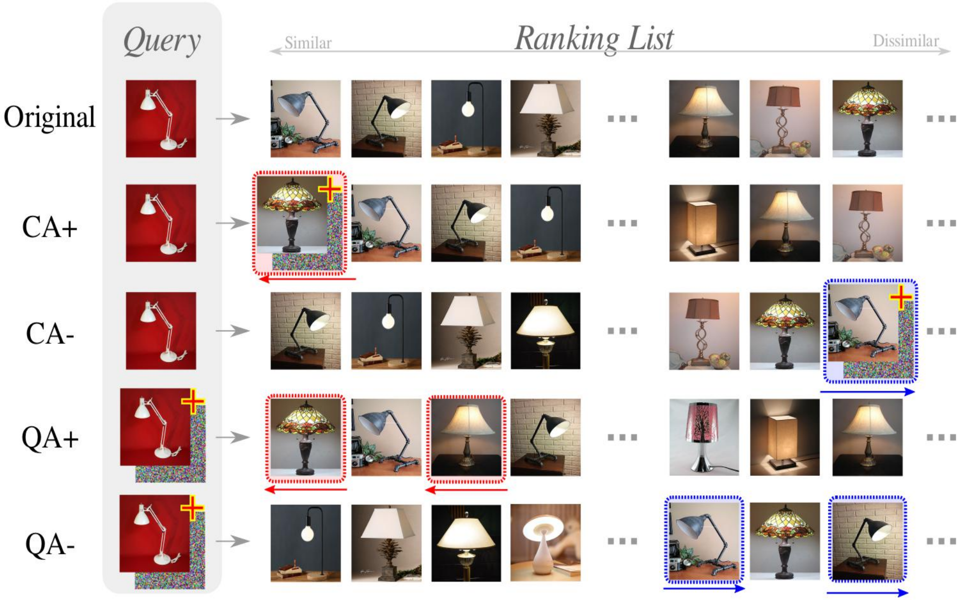
- Query Attack (QA): Raise (QA+) or lower (QA-) the rank by perturbing <u>queries</u>.



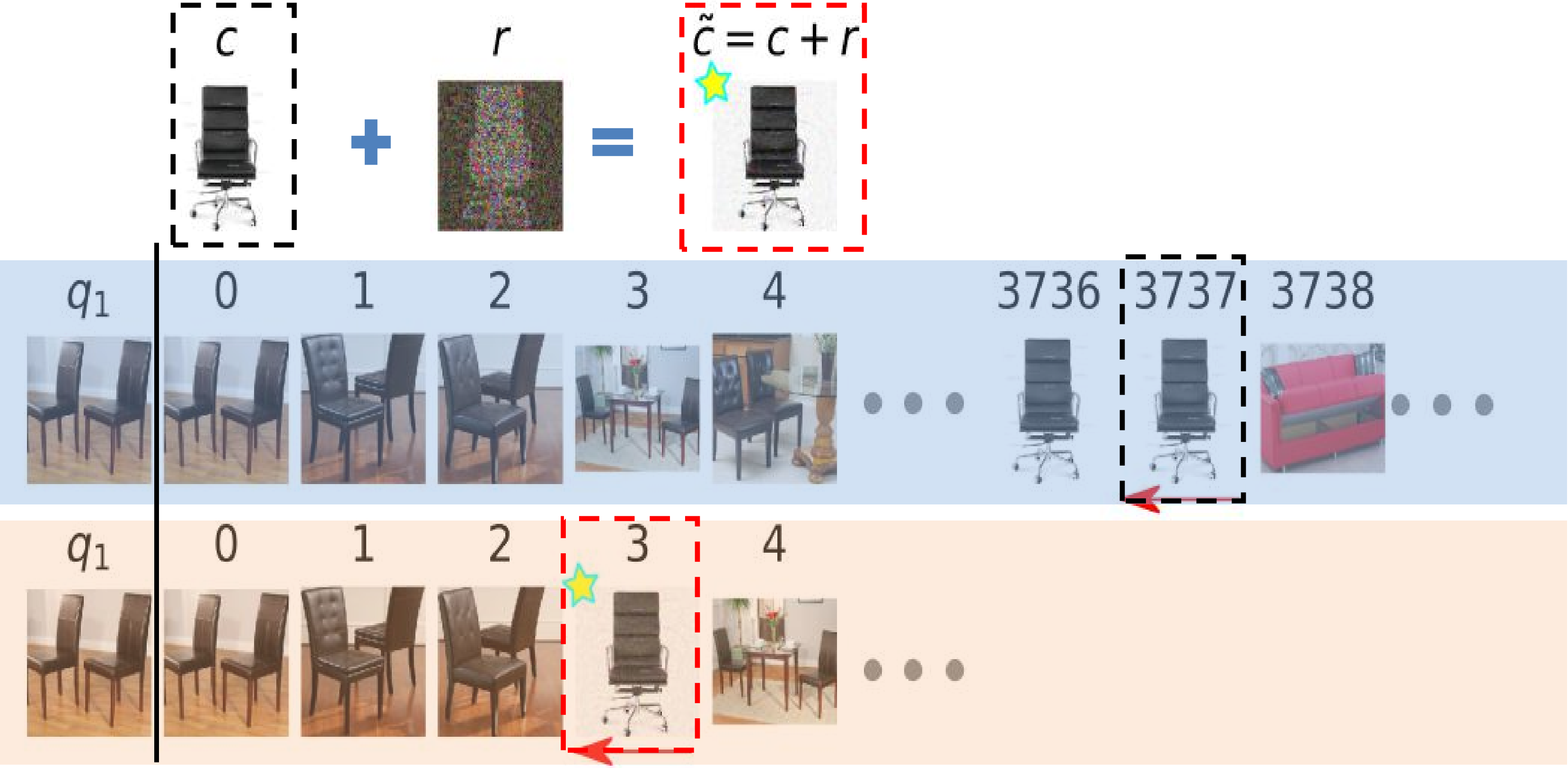
Adversarial Ranking Attack

Case1: a malicious seller may attempt to raise the rank of his/her own product (CA+), or lower the rank of his competitor's product (CA-);

Case2: a "man-in-the-middle" attacker (e.g., a malicious advertising company) could hijack the query image in order to promote (QA+) or impede (QA-) the sales of specific products.



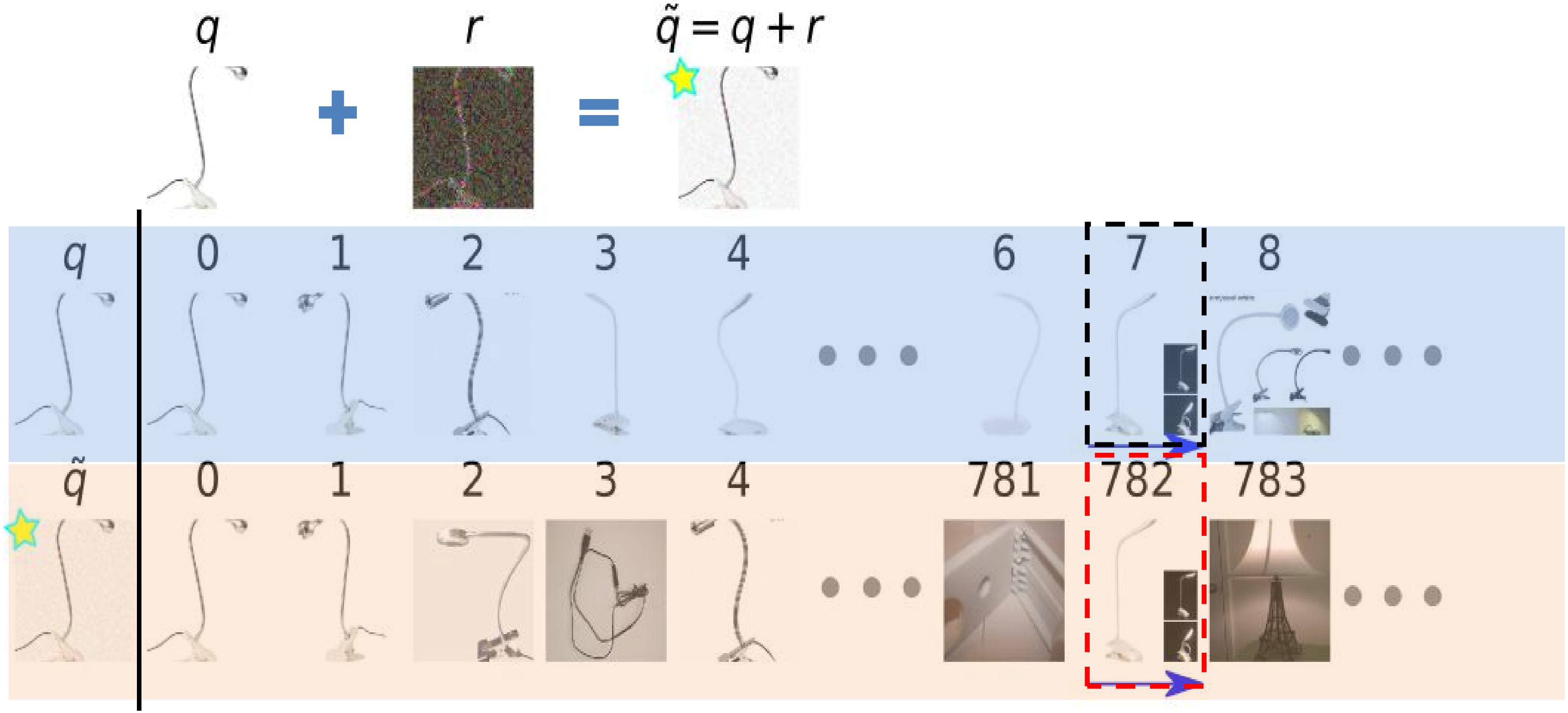
Showcase: CA+ Attack



Original ranking results

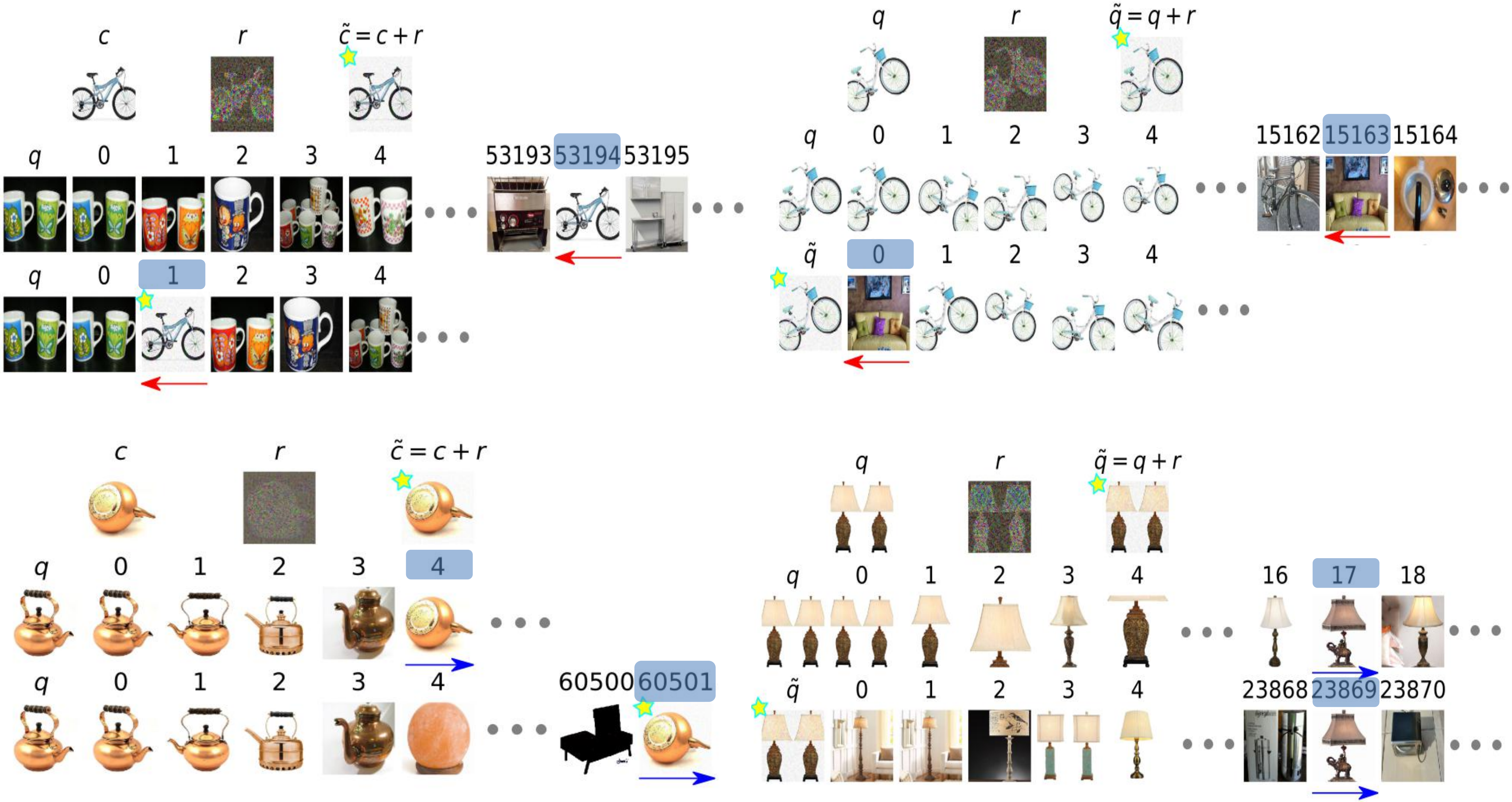
Ranking results after QA-

Showcase: QA- Attack



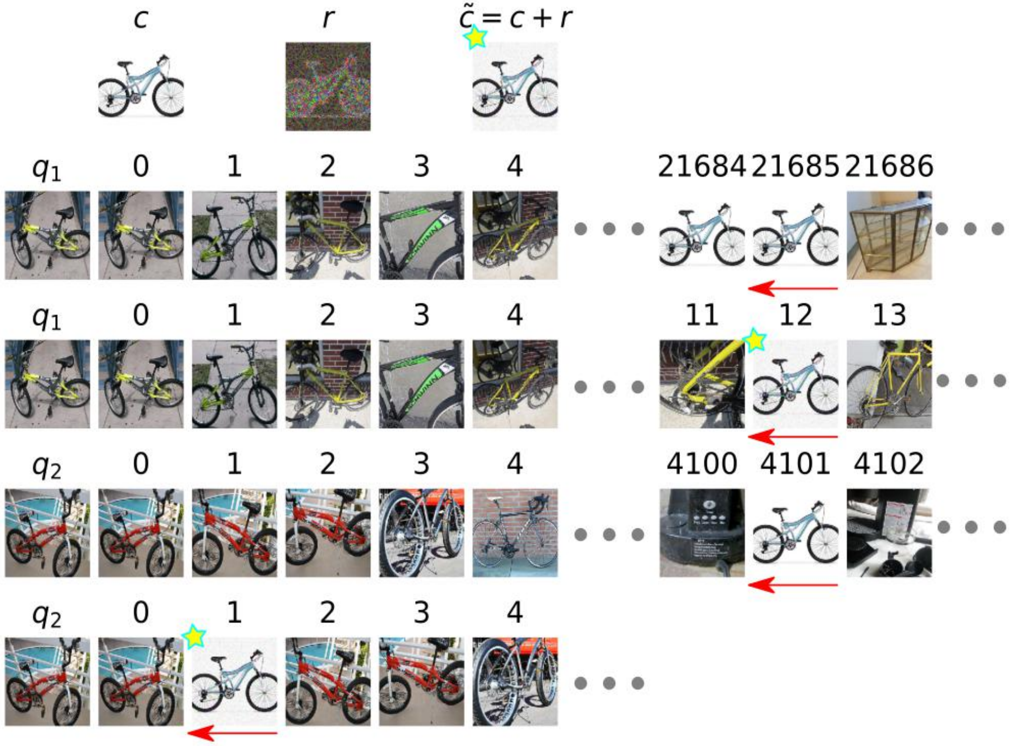
Original ranking order

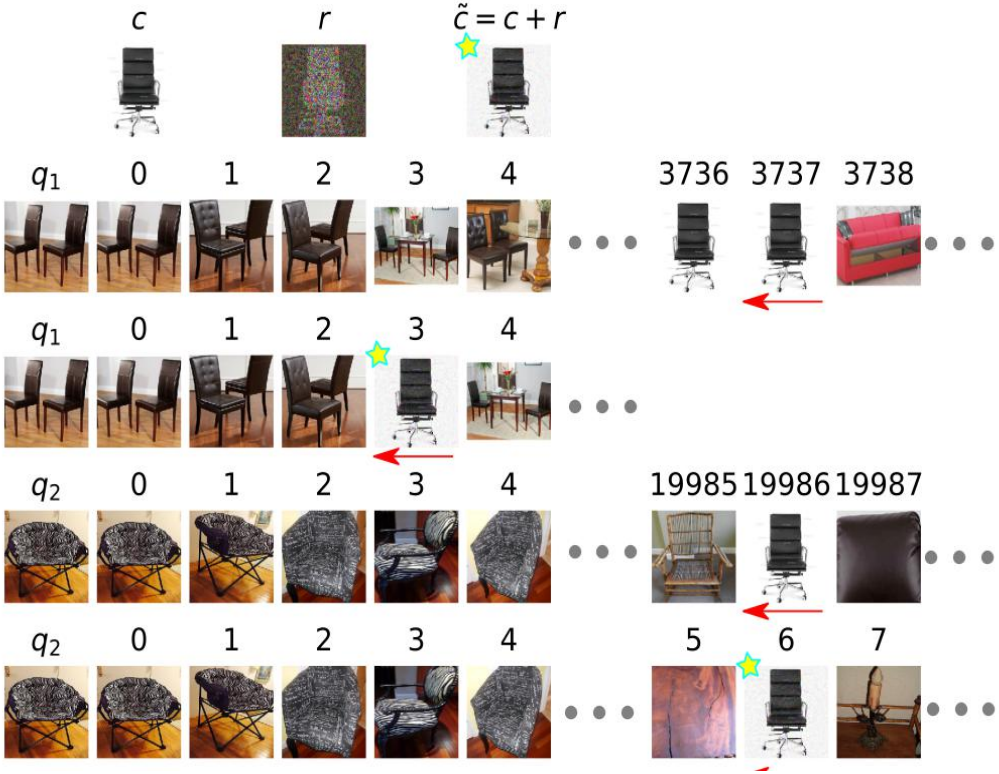
Ranking order after QA-



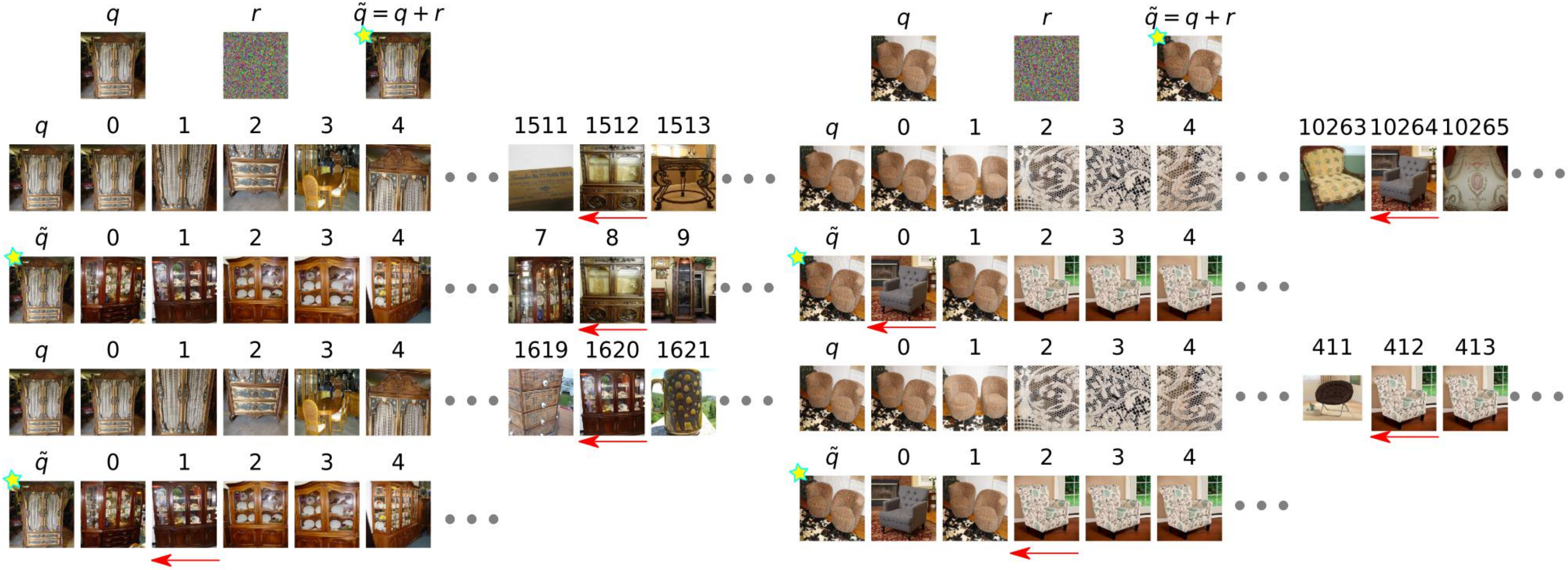
c
r
c̃ = c + r
q 0 1 2 3 4
53193 53194 53195
q 0 1 2 3 4
q
r
q̃ = q + r
q 0 1 2 3 4
15162 15163 15164
q̃ 0 1 2 3 4
c
r
c̃ = c + r
q 0 1 2 3 4
q 0 1 2 3 4
60500 60501
q
r
q̃ = q + r
q 0 1 2 3 4
16 17 18
q̃ 0 1 2 3 4
23868 23869 23870

Showcase: CA+ with a Query Set




c
r
$\tilde{c} = c + r$
q_1 0 1 2 3 4 3736 3737 3738
q_1 0 1 2 3 4
q_2 0 1 2 3 4 19985 19986 19987
q_2 0 1 2 3 4 5 6 7

Showcase: QA+ with a Query Set



Thanks!

c
r
$\tilde{c} = c + r$
q_1 0 1 2 3 4 3475 3476 3477
q_1 0 1 2 3 4 14 15 16
q_2 0 1 2 3 4 414 415 416
q_2 0 1 2 3 4 26 27 28
q_3 0 1 2 3 4 11639 11640 11641
q_3 0 1 2 3 4 8 9 10
q_4 0 1 2 3 4 11750 11751 11752
q_4 0 1 2 3 4
q_5 0 1 2 3 4 5069 5070 5071
q_5 0 1 2 3 4 119 120 121
q_6 0 1 2 3 4 5274 5275 5276
q_6 0 1 2 3 4 8 9 10
q_7 0 1 2 3 4 27232 27233 27234
q_7 0 1 2 3 4 22 23 24
q_8 0 1 2 3 4 1169 1170 1171
q_8 0 1 2 3 4 6 7 8
q_9 0 1 2 3 4 2369 2370 2371
q_9 0 1 2 3 4 68 69 70
q_{10} 0 1 2 3 4 15168 15169 15170
q_{10} 0 1 2 3 4 34 35 36

https://arxiv.org/abs/2002.11293