

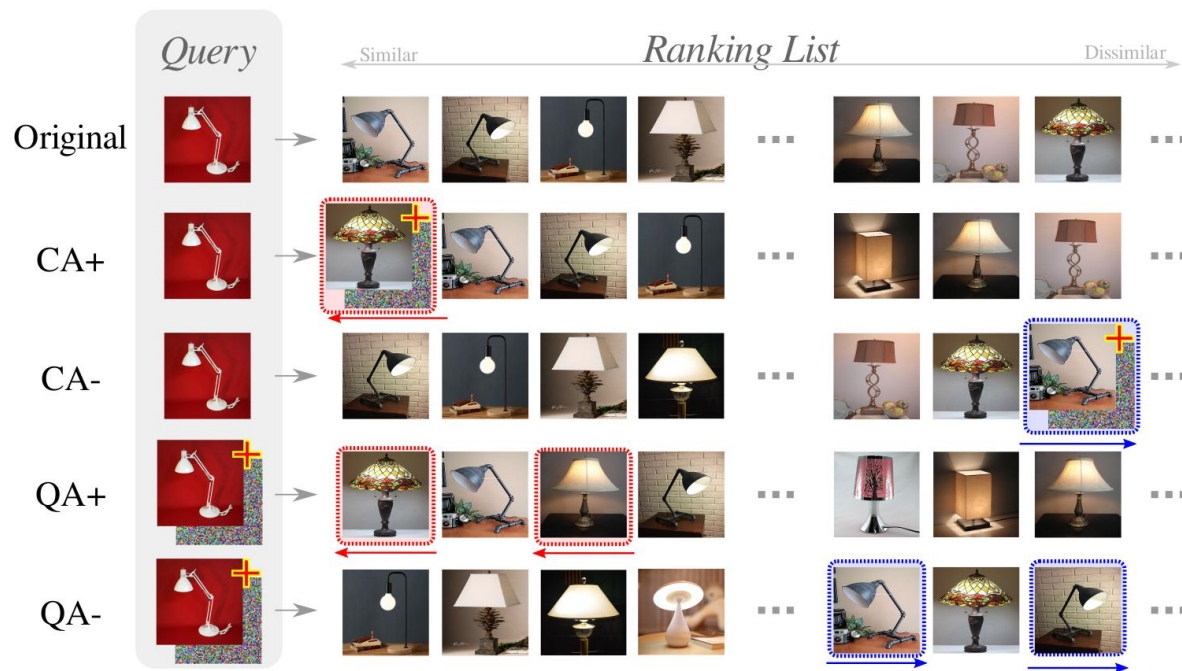
Adversarial Ranking Attack and Defense

Mo Zhou, Zhenxing Niu, Le Wang, Qilin Zhang, Gang Hua
May 2020

<https://arxiv.org/abs/2002.11293>

Adversarial Ranking Attack

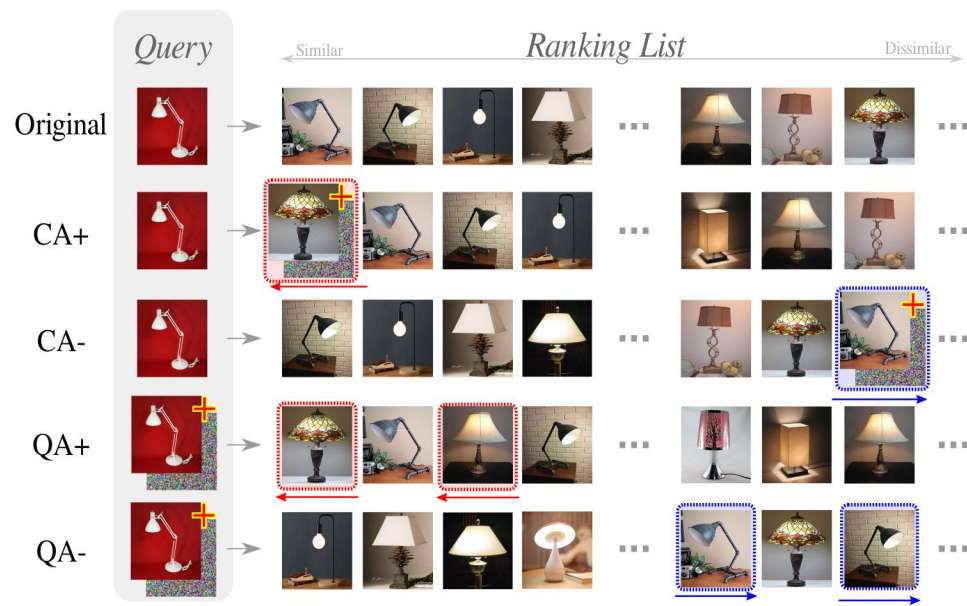
- **Definition:** **Raise** or **lower** the rank of chosen candidates with respect to a specific query set
- **Candidate Attack (CA):** Raise (CA+) or lower (CA-) the rank by perturbing candidates.
- **Query Attack (QA):** Raise (QA+) or lower (QA-) the rank by perturbing queries.



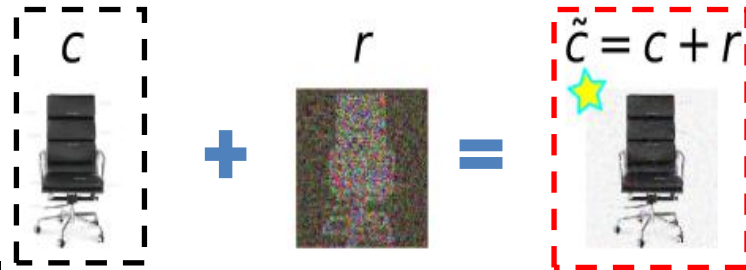
Adversarial Ranking Attack

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-);

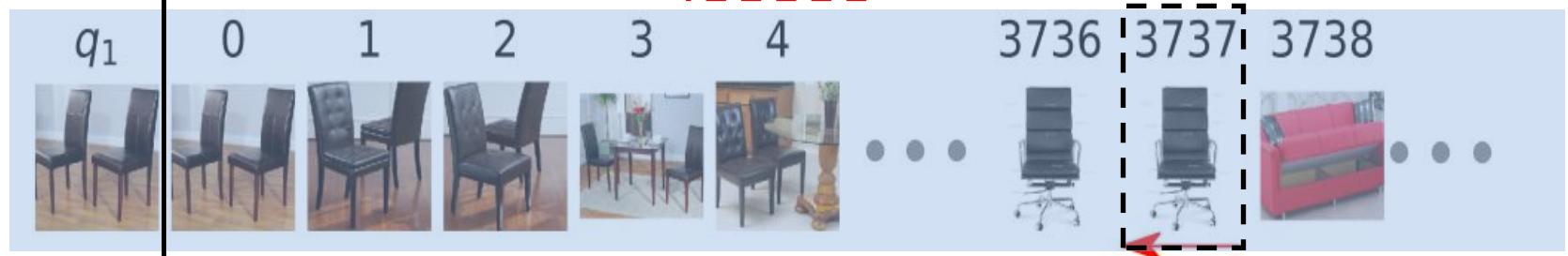
Case2: a “man-in-the-middle” attacker (e.g., a malicious advertising company) could **hijack** the query image in order to promote (QA+) or impede (QA-) the sales of specific products.



Showcase: CA+ Attack

$$\begin{array}{|c|} \hline c \\ \hline \end{array} + \begin{array}{|c|} \hline r \\ \hline \end{array} = \begin{array}{|c|} \hline \tilde{c} = c + r \\ \hline \end{array}$$


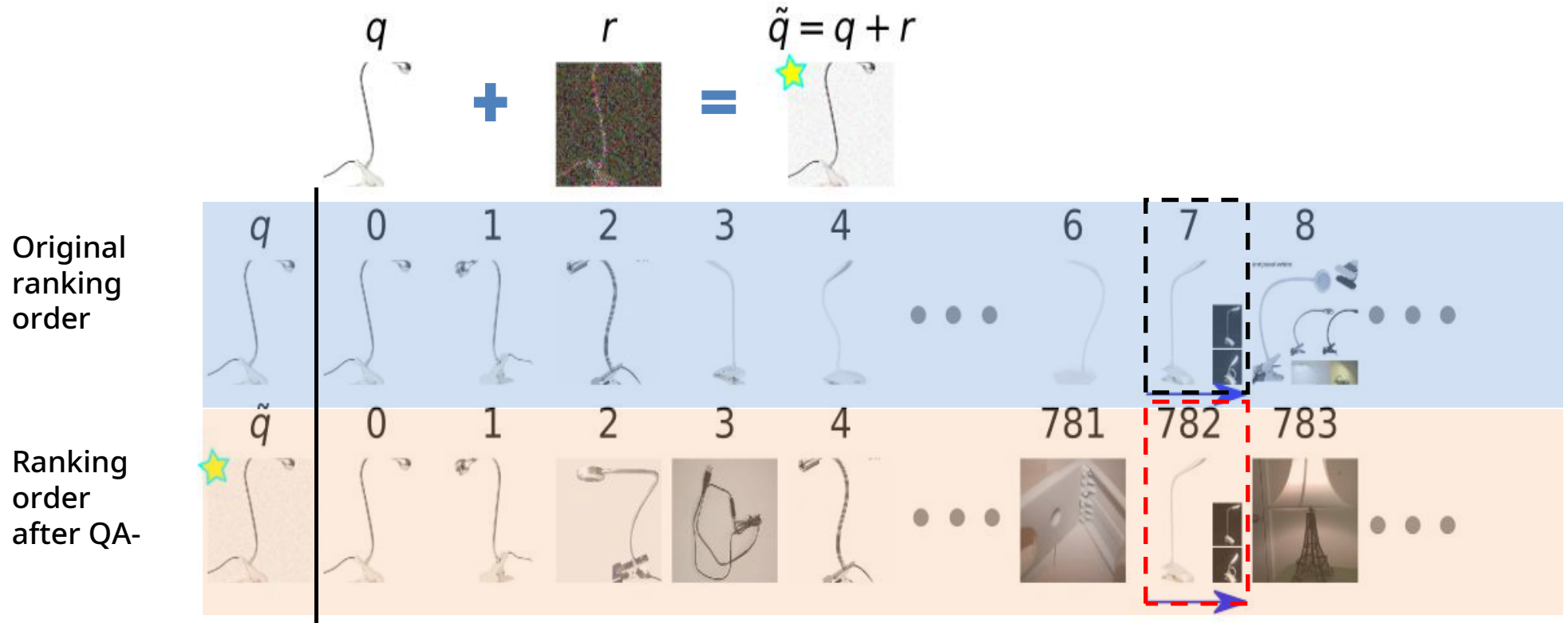
Original
ranking
results

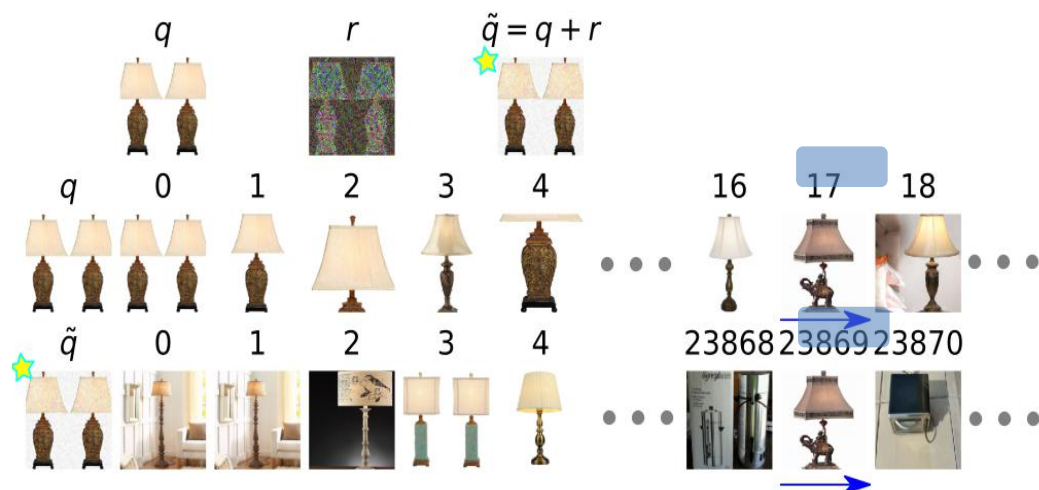
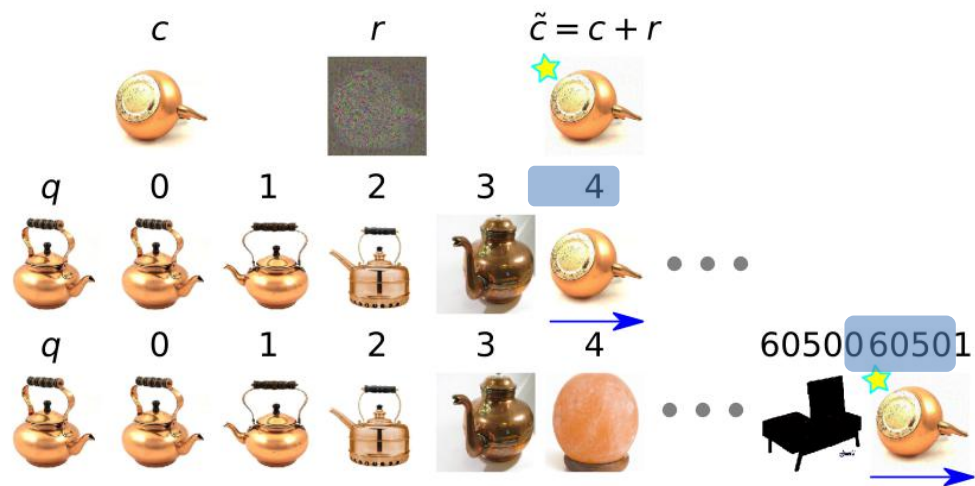
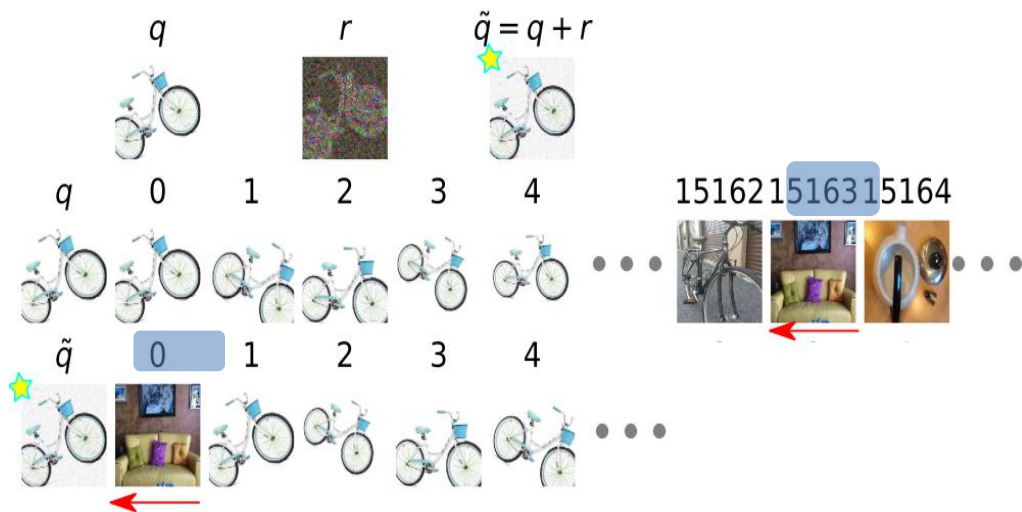
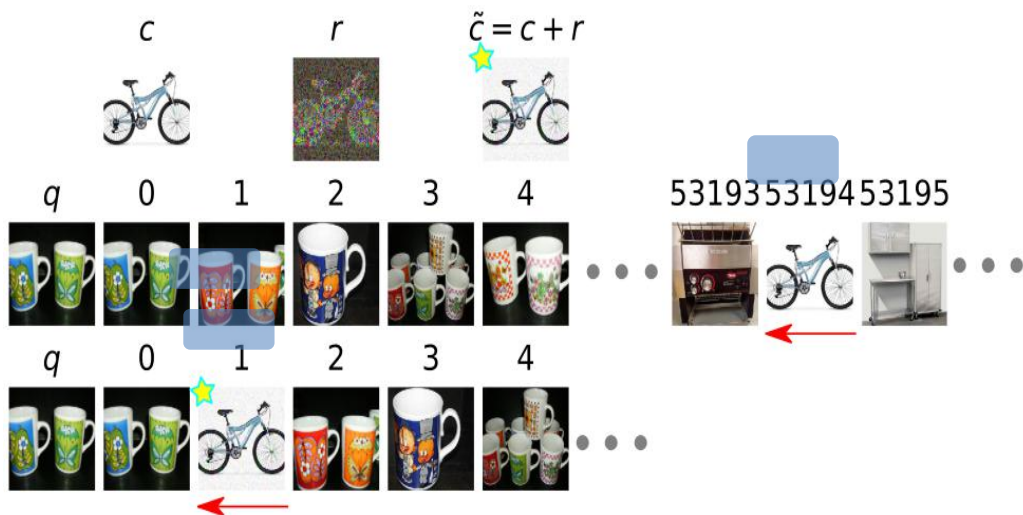


Ranking
results
after QA-

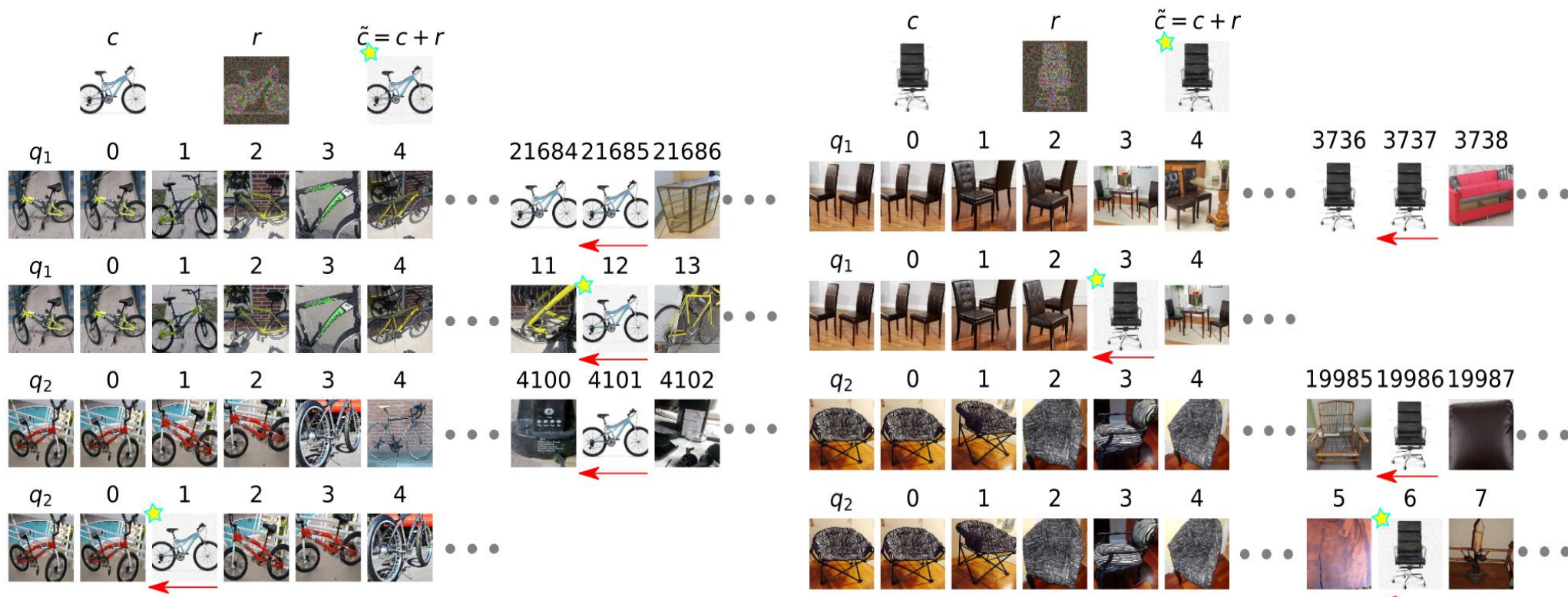


Showcase: QA- Attack

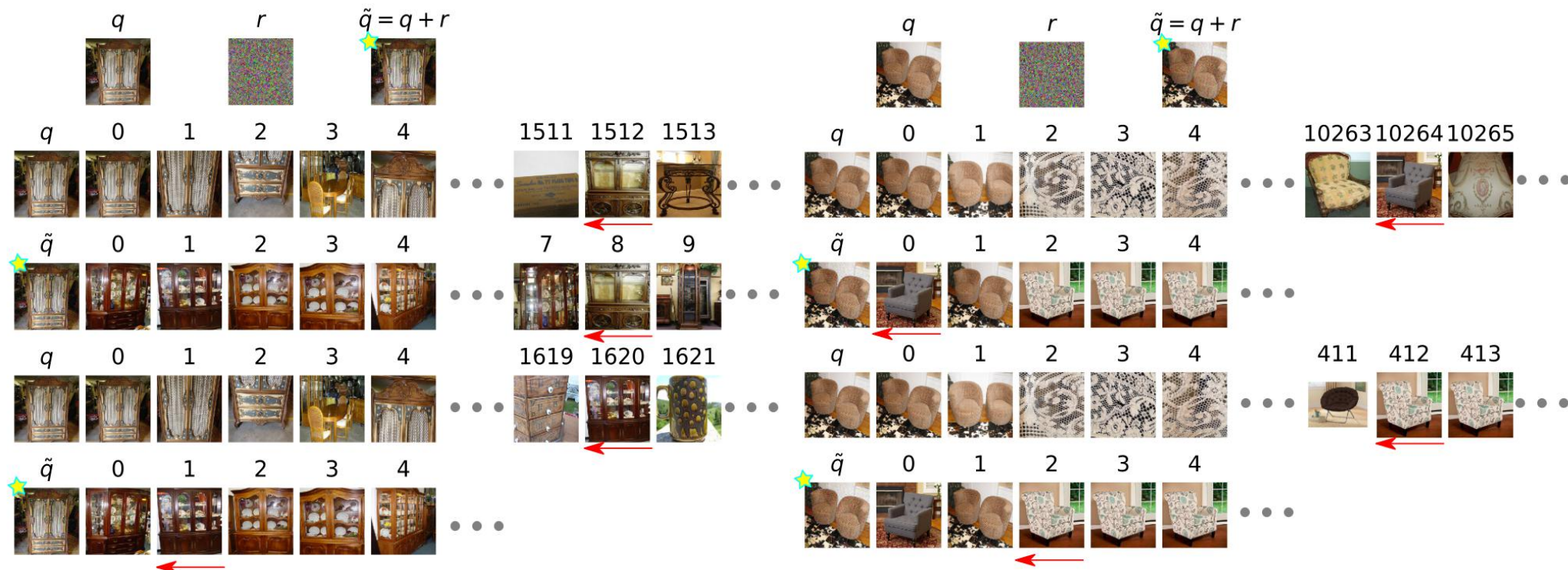




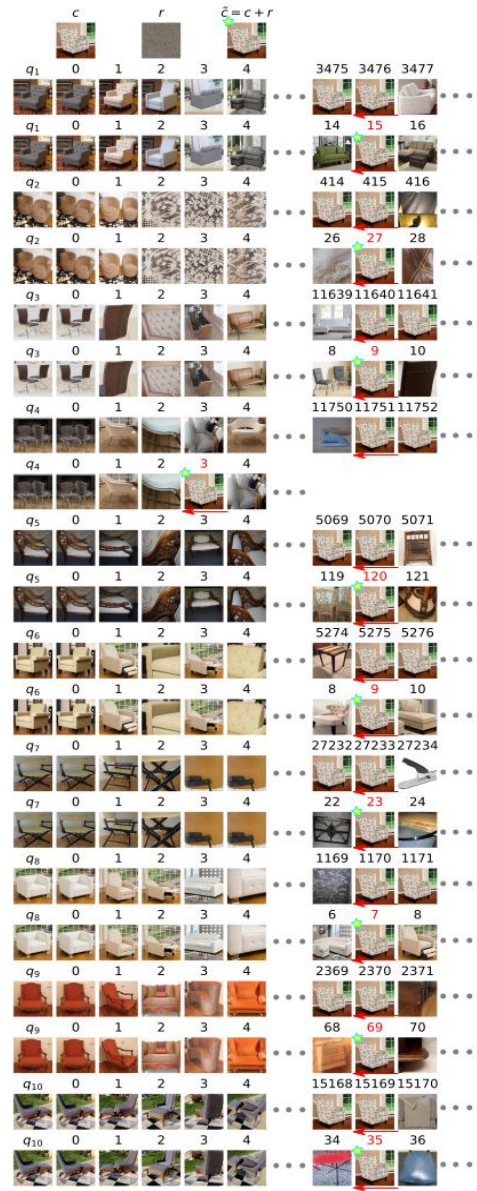
Showcase: CA+ with a Query Set



Showcase: QA+ with a Query Set



Thanks!



<https://arxiv.org/abs/2002.11293>