Learning to Attack Real-World Models for Person Re-Identification via Virtual-Guided Meta-Learning

Fengxiang Yang¹, Zhun Zhong², Hong Liu³, Zheng Wang³, Zhiming Luo¹, Shaozi Li¹, Nicu Sebe¹,², Shin’ichi Satoh³
¹Xiamen University ²University of Trento ³National Institute of Informatics ⁴University of Tokyo ⁵Huawei Research, Ireland

Motivation: Verifying and improving the robustness of re-ID models.
Approach: Learning adversarial examples to corrupt re-ID models.

Challenges

- Domain shift and open-set property in re-ID requires attackers to adapt to different environments, i.e., attackers should be universal.
- Recent works[1,2] on re-ID attack generate adversarial examples individually and are not efficient enough.

Our Solution & Contributions

How to achieve efficient?: Universal Adversarial Perturbation[3]. Why UAP?: Simplify the attack by adding UAP to queries.
How to achieve universal?: Synthetic Data & Meta-Learning. Why Virtual Data?
(1) Easy to collect (2) Privacy-free (3) Balanced data distribution.

Contributions:
(1) Meta-learning strategy. (2) Virtual data for optimization. (3) Inspiration of improving robustness obtained from visualization.

Visualization of Robust Queries

We visualize robust queries that survived from our attack and have 2 findings:
Finding 1: Occlusion is robust to attack. Suggestion 1: Erasing may improve robustness[4].
Finding 2: Camera styles are robust. Suggestion 2: Camera styles may improve robustness.

Motivation: Verifying and improving the robustness of re-ID models.
Approach: Learning adversarial examples to corrupt re-ID models.

Contact Us
If you have any problem, please send email to us (yangfx@stu.xmu.edu.cn) or ask in Github.
Scan the right QR code for code and other resources.

References