

Person Re-identification and Recognition: Sachin and Hem (AI intern 2021)

Overview:

- The main objective of person reidentification is to find a probe person from whole scene which shows great significance in video surveillance field to track lost people, re-identification, and verification of person.
- The typical person reidentification completes pedestrian detection, feature extraction, Photometric and geometric transformation and similarity estimation.

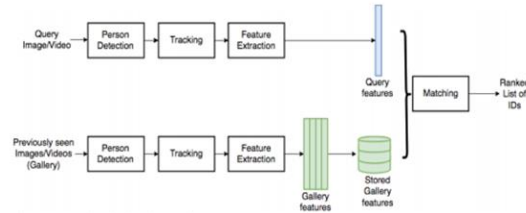


Figure 1: Typical pipeline of a re-ID system

- Person Re-ID and person search are extremely challenging due to human posture, light, video camera position, low resolution, occlusion, viewpoints, Uniform clothing, Unreliable bounding box generation, pose estimation and different variations in images.
- Person (re-ID) is applied in academic community and large-scale industry implementation, such as public safety, tracking of person in widespread public parks, universities and streets, behavior analysis and surveillance.

Person Re-Identification with Deep Learning: Process

- **Collection** of raw video dataset
- **Bounding box** generation uses the extraction of person images from video data - Training data annotation is usually done for discriminative re-identification model learning due to the large cross-camera variations
- Model training:** Development phase of re-identification system where the system handles the various challenges, color calibration, transfer learning, descriptor learning, distance metric learning, or their combinations

- **Pedestrian retrieval:** Testing phase where re-identification model extracts feature from a given query for a gallery-set. A ranking list is obtained by sorting the calculated query to gallery similarity i.e., probability of ID-match

State-of-the-Art Re-Identification: Closed-World: A standard closed-world Re-ID system contains three main components:
1.Feature Representation Learning: Focuses on developing the feature construction strategies, Four different feature learning strategies.

- Global Feature:** learning a global representation for each person image
 - Local Feature:** learning part-aggregated local features
 - Auxiliary Feature:** learning the feature representation using auxiliary information
 - Video Feature:** learning the video representation using multiple image frames and temporal information
- 2.Deep Metric Learning:** designing the training objectives with different loss functions or sampling strategies; Types of loss:

I. Identity Loss.

- Treats the training process of person Re-ID as an image classification problem.
- To avoid the model fitting to over-confident annotated labels, improving the generalizability

II. Triple loss: Treats the Re-ID model training process as a retrieval ranking problem.

: Distance between the positive pair should be smaller than the negative pair by a pre-defined margin.

III. Verification loss: Optimizes the pairwise relationship, either with a contrastive loss, or binary verification

3.Ranking Optimization: To utilize the gallery-to gallery similarity to optimize the initial ranking list.

- Query Adaptive strategy replaces the uniform searching engine to improve the performance.
- Human Interaction involves using human feedback to optimize the ranking list.

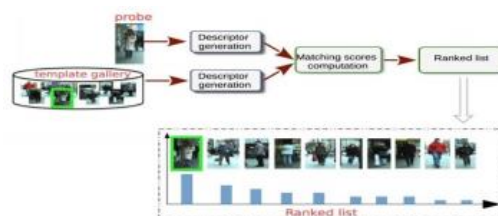


Fig2: Re-Identification: Closed-World

Open-World Person Re-Identification

- Heterogeneous Re-ID** by matching person images across heterogeneous modalities
- End-to-End Re-ID** from raw videos and images, the reliance on the additional step for bounding boxes generation.
- Semi-supervised and Unsupervised Re-ID** with limited or unavailable annotated labels.
- Noise-Robust Re-ID** which includes partial Re-ID with heavy occlusion, Re-ID with sample noise caused by detection or tracking errors and Re-ID with label noise caused by annotation error.
- Open-set Re-ID and Beyond** is usually formulated as a person verification problem, such as discriminating whether two person-images belong to the same identity