Aligned Re-id

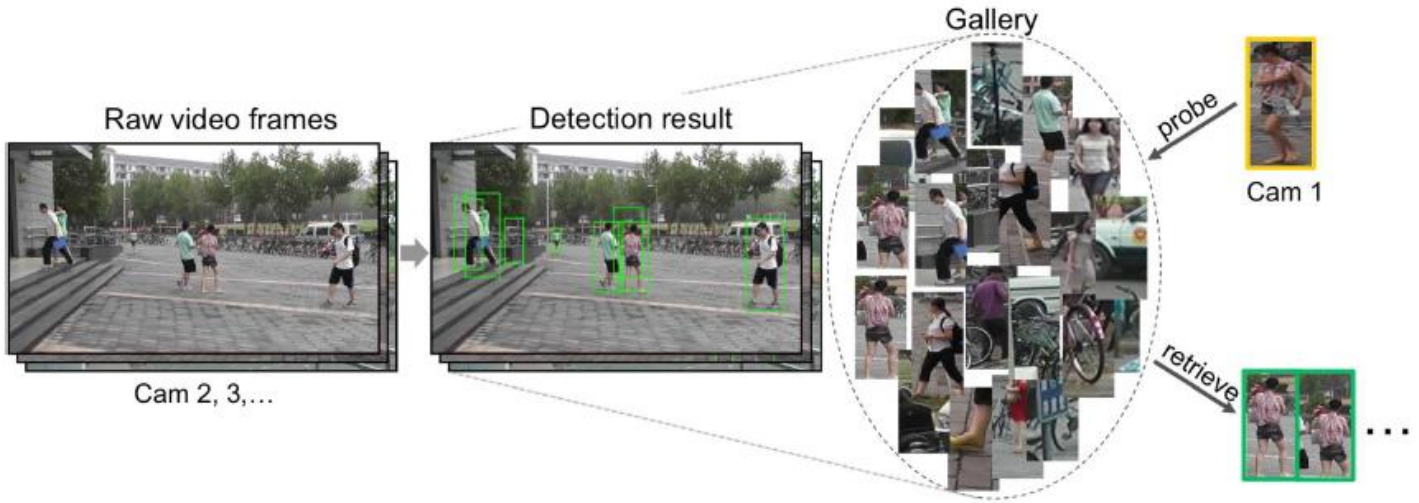
Ziqiang Zheng

What’s Re-id

Introduction

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Person Detection **Person retrieval / re-identification**

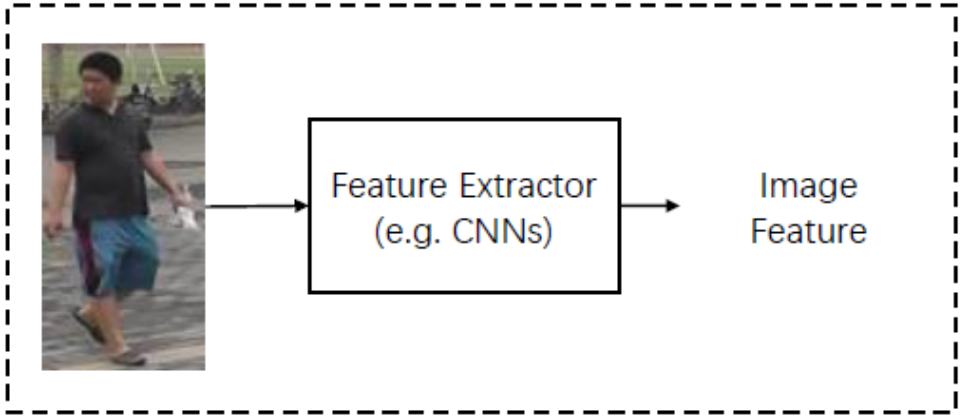


Person retrieval / re-identification

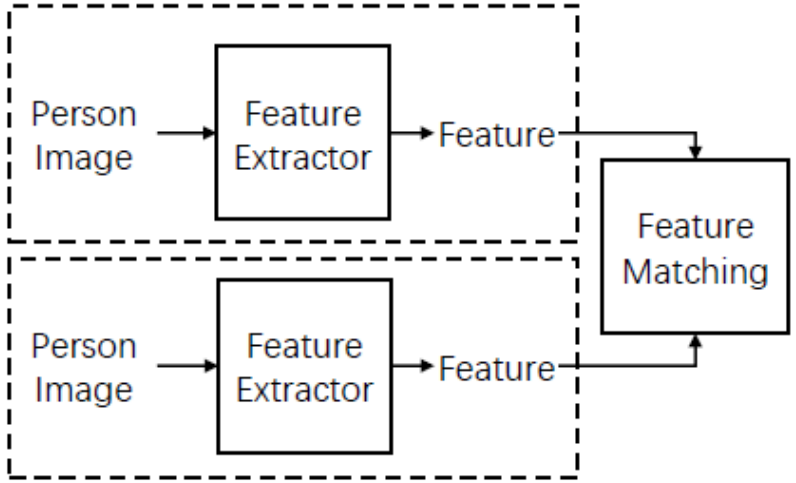
General methods

- Step 1: Feature Extraction
  - Extracting features for every person images
- Step 2: Feature Matching
  - Matching features to calculate the similarity score

From 2014, deep models were used to improve these two parts



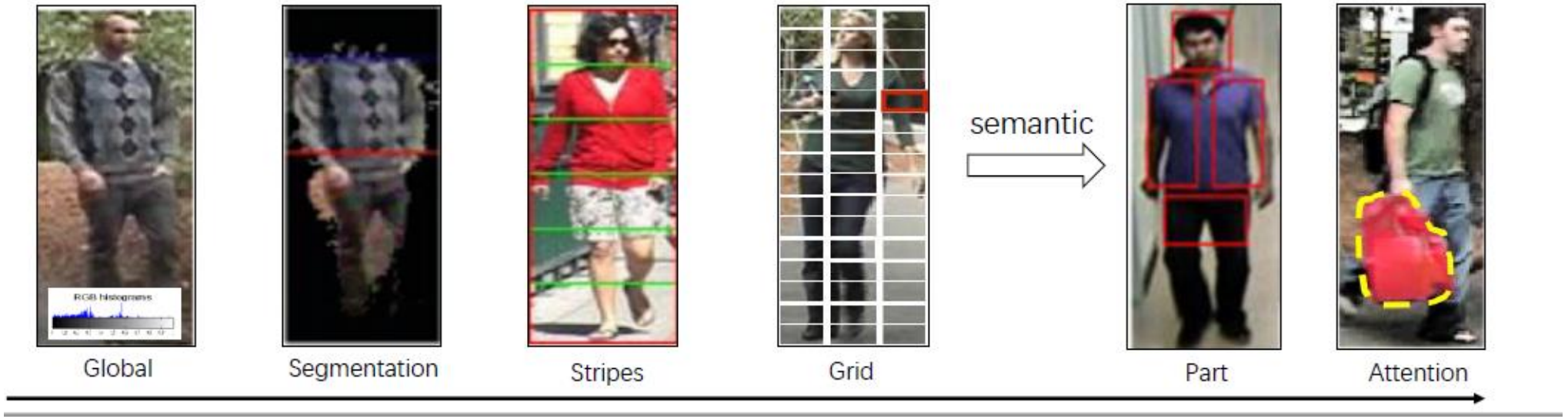
Feature Extraction



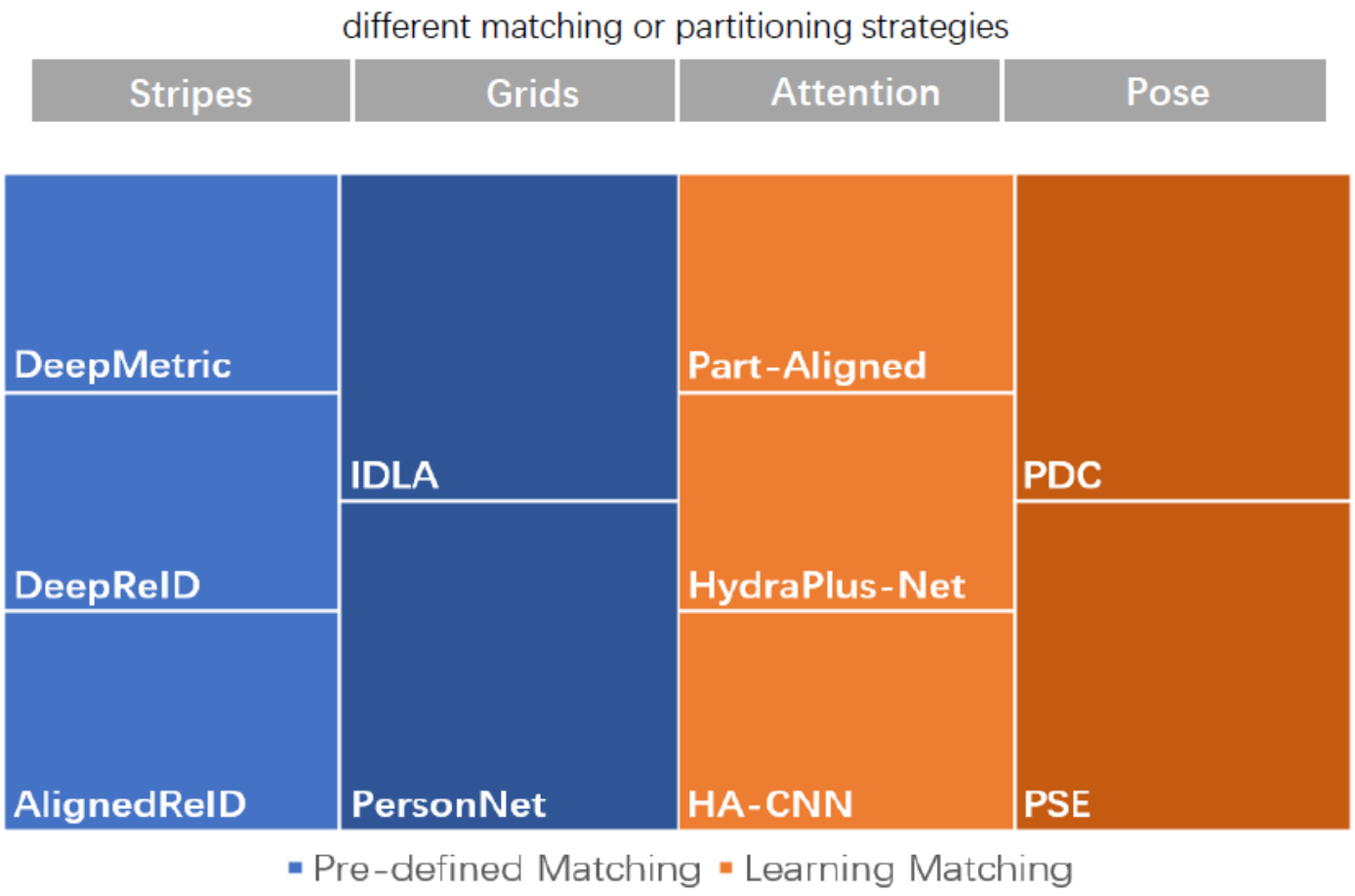
Feature Matching

Feature Matching Methods

- Matching based on pre-defined locations
  - Global, local stripes, grid patches
- Matching based on semantic regions
  - Person parts, salient regions, attention regions



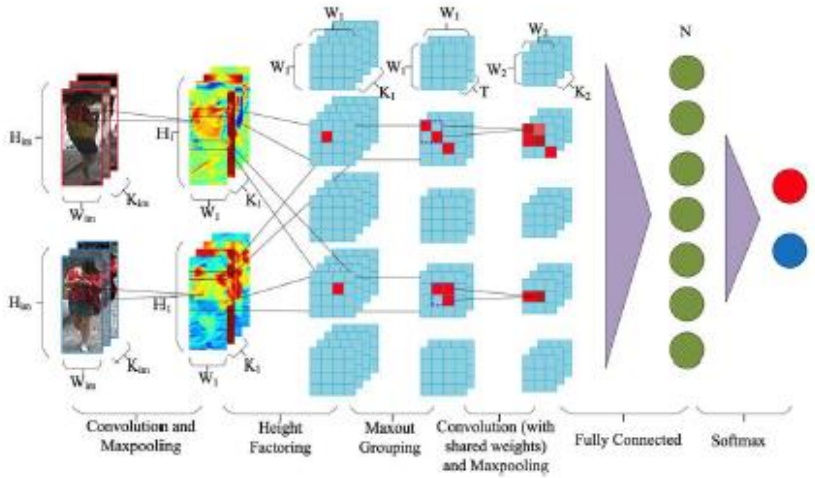
Deep learning methods



CNN-based methods

- Before 2015, researchers addressed this task using an expensive matching procedure.

Zhao et al. CVPR 2013



Li et al. CVPR 2014

Two images are first compared using their local regions. The local similarities are aggregated into a global similarity.

Some universal methods

- （1）leverage external cues（human pose estimation）
- （2）abandon cues from semantic parts
- （3）Global feature and local feature
- （4）Pose alignment/matching

Drawback

Figure 1. Challenges in ReID: (a-b) inaccurate detection, (c-d) pose misalignments, (e-f) occlusions, (g-h) very similar appearance.

Solution

- a global feature and local features jointly
- mutual learning approach
- Metric Learning distance metric learning
- two branches

Calculating the shortest path (DP)

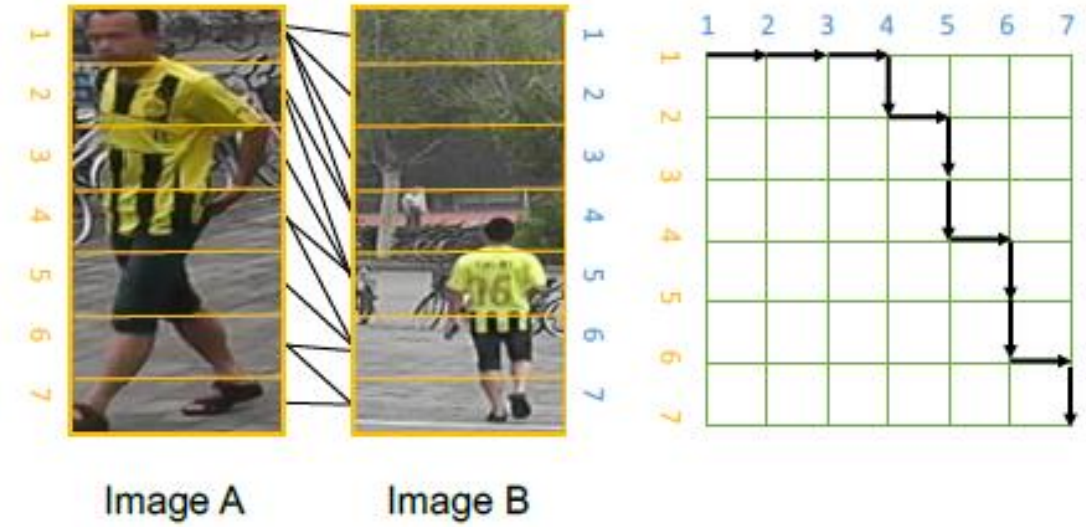


Figure 3. Example of AlignedReID local distance computed by finding the shortest path. The black arrows show the shortest path in the corresponding distance matrix on the right. The black lines show the corresponding alignment between the two images on the left.

Two branches

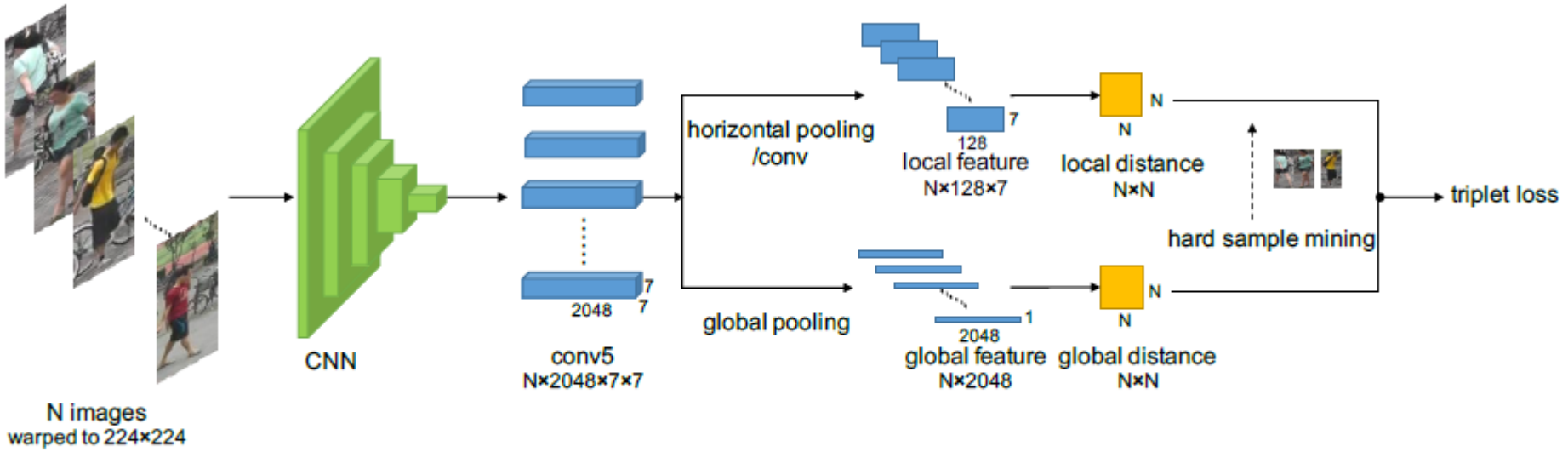


Figure 2. The framework of AlignedReID. Both the global branch and the local branch share the same convolution network to extract the feature map. The global feature is extracted by applying global pooling directly on the feature map. For the local branch, one 1×1 convolution layer is applied after horizontal pooling, which is a global pooling with a horizontal orientation. Triplet hard loss is applied, which selects triplet samples by hard sample mining according to global distances.

GAP (global average pooling)

- Attention map
- Replace FC