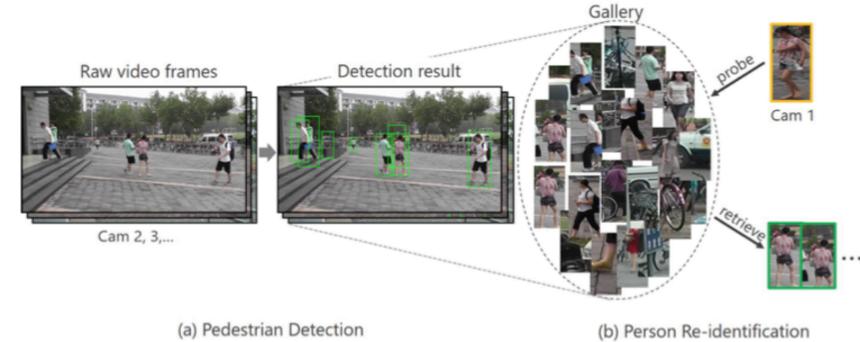


Introduction

Previous works: person re-ID only

Our work: person detection + re-ID



Contribution

We introduce the PRW dataset

- We evaluate person detection and re-ID as a whole
- Input: raw video frames, a query
- Output: a rank list of bounding boxes
- A gallery has to be generated using a particular detector
- Detector errors may lead to decrease in re-ID accuracy
- The detector threshold determines the gallery size
- uses mAP & CMC under a fixed number of bboxes per image, e.g., 3 or 5 bounding boxes per frame

We propose ID-discriminative Embedding (IDE)

- easy to train/test
 - produces competitive accuracy on iLIDS-VID and PRID-2011 (our ECCV16 paper).
- Insights on how detection helps re-identification**
- Evaluate detector performance under re-ID application
 - A cascade IDE fine-tuning strategy: first fine-tune detection, then fine-tune re-ID
 - Confidence Weighted Similarity: integrating detection confidence in re-ID matching scores

PRW (Person Re-identification in the Wild)

Statistics

Datasets	#frame	#ID	#anno. Box	#box/ID	#gallery box	#cam
PRW	11,816	932	34,304	36.8	100-500k	6
CAMPUS	214	74	1,519	20.5	2k-4k	3
EPFL	80	30	294	9.8	~1k	4
Market	0	1,501	25,259	19.9	19,732	6
RAiD	0	43	6,920	160.9	6,920	4
VIPeR	0	632	1,264	2	1,264	2
i-LIDS	0	119	476	2	476	2
CUHK03	0	1,360	13,164	9.7	13,164	2

Annotation procedure



- All appearing pedestrians are assigned an ID
- If we are not sure about a person's ID, we assign -2 to it.

- ✓ Different detectors create galleries different in size, bbox qualities, etc.
- ✓ One needs to build his own gallery
- ✓ Generally, a better detector leads to higher re-ID results, but how to define "better"?
- ✓ Detector errors will propagate to re-ID

A gallery produced by DPM

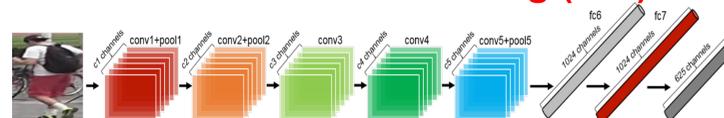


Evaluation and Baseline

Evaluation Details

- ✓ Cross-camera search.
- ✓ Fixed train/test partition
- ✓ Train: 482 IDs, 5,704 frames
- ✓ Gallery: 450 IDs, 6,112 frames
- ✓ 2,057 queries, 4,57 queries per ID
- ✓ Gallery size depends on the detector and the detection threshold
- ✓ Evaluation: mAP and rank-n accuracy under a fixed number (e.g., 3 or 5) of bboxes per frame

ID-discriminative Embedding (IDE)



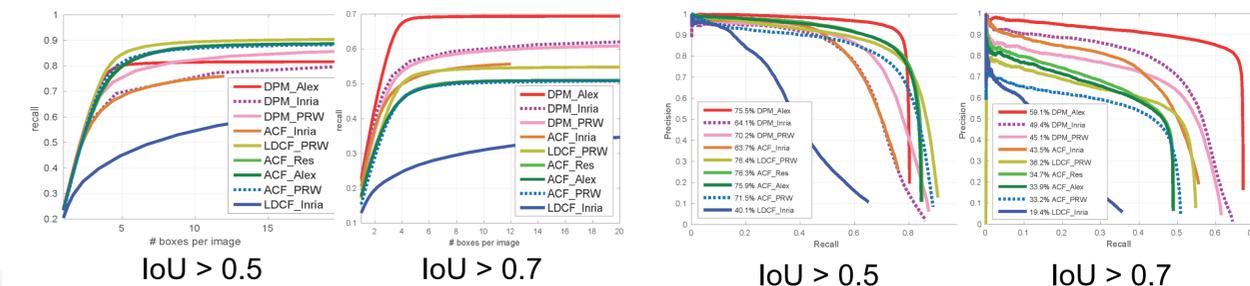
- ✓ Train a CaffeNet on the training set
- ✓ Classify each training bbox into one of the 482 IDs
- ✓ We extract FC7 for each bbox during testing
- ✓ Euclidean distance is used for similarity

We call for attention to a contemporary work: Xiao et al. Joint Detection and Identification Feature Learning for Person Search, CVPR 2017.

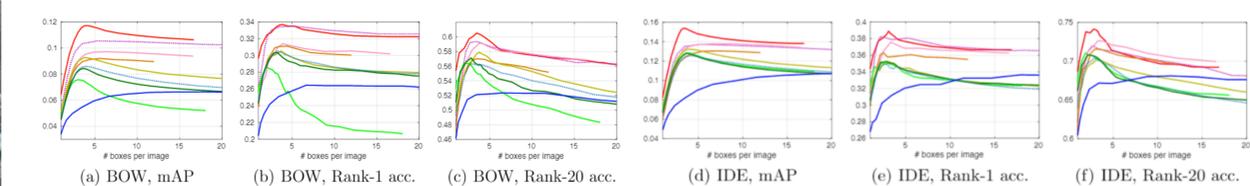
Empirical Studies

How to evaluate person detection under the application of re-ID?

The accuracy of various detectors on PRW, under IoU > 0.5 or 0.7

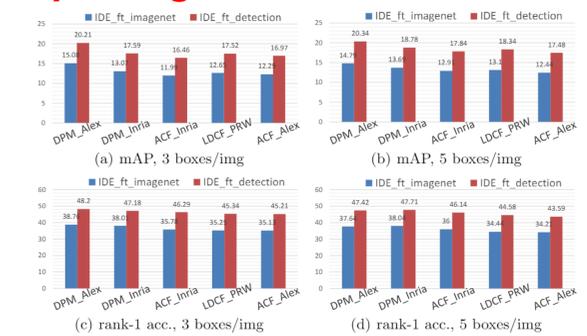
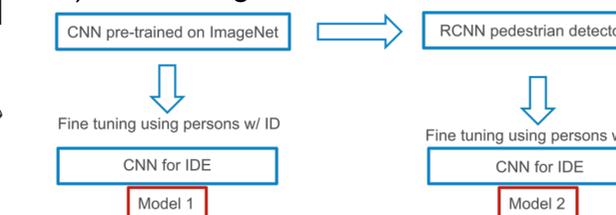


Re-ID accuracy under different detectors and different numbers of bboxes per frame



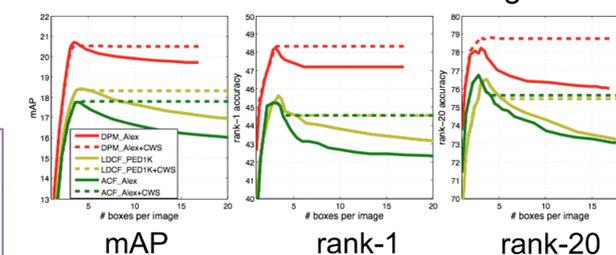
A cascade fine-tuning strategy for improving IDE

- 1) Fine-tuning 2-class CNN: person and background
- 2) Fine-tuning 482-class CNN, i.e., IDE



Confidence Weighted Similarity

We multiply the normalized detection confidence into the re-ID matching scores



Sample re-ID results on PRW

