Keypoint Promptable Re-Identification

Vladimir Somers [1,2,3] Alexandre Alahi [2] Christophe De Vleeschouwer [1]

[1]UCLouvain [2]EPFL [3]Sportradar

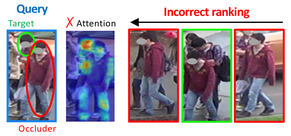
EUROPEAN CONFERENCE ON COMPUTER VISION

Overview

We present **KPR** (**K**eypoint **P**romptable **R**eID), a novel re-identification method that can be prompted with **positive and negative keypoints** and that is robust to **multi-person occlusions**.

Motivations

We aim to solve **Multi-Person Ambiguity** (**MPA**), an important issue overlooked by previous occluded ReID methods. **MPA** arises when multiple individuals are visible in the same bounding box, making it challenging to determine the **intended ReID target**.



(a) Existing methods (e.g. *BPBreID*) are confused by the Multi-Person Ambiguity (**MPA**)

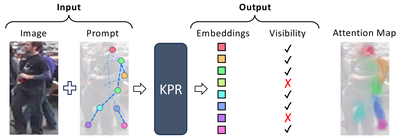
(b) KPR explicitly addresses **MPA** with positive (pastel) and negative (red) keypoints prompt

The Occluded-PoseTrack-ReID dataset

To support research on promptable reID, we release **Occluded-PoseTrack-ReID**, a multi-person **occluded** ReID dataset with high-quality **keypoints annotations**, that is derived from the PoseTrack21 dataset for pose tracking. We also release **new keypoint annotations** for four popular ReID datasets.

KPR: Keypoint Promptable Re-Identification

KPR takes an image with **optional positive/negative** keypoints prompts as input and produces **part-based features** of the prompted target together with their **visibility scores**.



(a) Keypoint Promptable ReID (KPR)

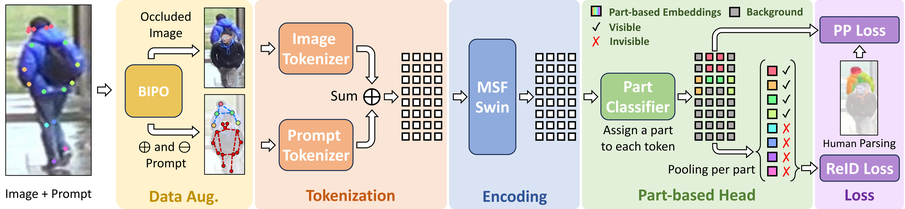
(b) Similarity = average distance between visible part embeddings

(c) Back (occluded) person prompted

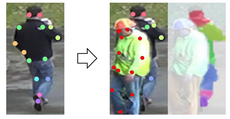
(d) Front person prompted

KPR Architecture

Our novel architecture includes: the **BIPO** data augmentation, a **Prompt Tokenizer**, Multi-stage Feature Fusion Swin-Transformer (**MSF**), and a **Part-based Head**.



BIPO Data Augmentation



Our proposed **B**atch-wise **I**nter-**P**erson **O**cclusion (**BIPO**) data augmentation creates artificial person occlusions that are consistent across image, prompt, and human parsing labels.

Comparison with SOTA and Ablations Studies

Experiments on **person retrieval** and multi-person **pose tracking**.

Datasets	Market-1501		Occluded-reID		Partial-reID		Occluded-PoseTrack	
Type	Holistic		Occluded					
Object Occlusions			✓		✓			
Person Occlusions							✓	
Methods	R-1	mAP	R-1	mAP	R-1	R-3	R-1	mAP
BoT [23]	94.5	85.9	58.4	52.3	-	-	*78.8*	*69.7*
PCB [33]	93.8	81.6	-	-	-	-	*81.7*	*71.2*
VGTri [40]	-	-	81.0	71.0	85.7	93.7[3]	-	-
PVPM [7]	-	-	66.8	59.5	78.3	-	-	-
HOReID [35]	94.2	84.9	80.3	70.2	85.3	91.0	-	-
ISP [47]	95.3	88.6	-	-	-	-	-	-
PAT [18]	95.4	88.0	81.6	72.1	88.0[3]	92.3	-	-
TRANS [11]	95.2	88.9	-	-	-	-	*83.5*	*73.4*
SOLIDER [4]	**96.9**[1]	**93.9**[1]	-	-	-	-	*84.4*	*76.1*[3]
SSGR [39]	96.1	89.3	78.5	72.9	-	-	-	-
FED [37]	95.0	86.3	**86.3**[1]	79.3[3]	84.6	-	-	-
BPBreid [31]	95.7	89.4	82.9	75.2	-	-	*84.9*	*75.5*
PFD [36]	95.5	89.7	83.0	81.5[2]	-	-	-	-
KPR_{IN} w/o prompt	95.6	88.7	83.3	78.2	81.7	86.0	85.3	75.4
KPR_{IN}	95.9	89.6	85.4[2]	79.1	86.0	90.0	**92.3**[1]	**82.3**[1]
KPR_{SOL} w/o prompt	96.6[2]	93.0[3]	80.0	78.5	90.3[2]	93.7[2]	86.1[3]	75.8
KPR_{SOL}	96.6[2]	93.2[2]	84.8[3]	**82.6**[1]	**90.7**[1]	**94.0**[1]	90.6[2]	81.2[2]

(a) Comparison with SOTA ReID methods

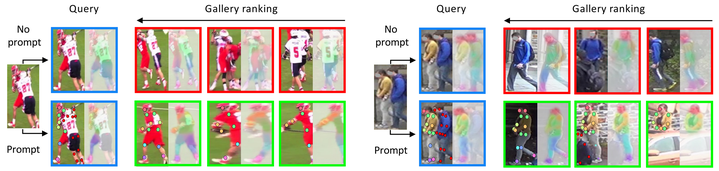
Method	HOTA↑	DetA↑	AssA↑	IDs↓
TRMOT [38]	46.8	40.9	55.0	-
FairMOT [43]	53.5	47.4	61.4	-
Tracktor++ [2]	58.3	52v.7	65.4	-
CorrTrack + ReID [5]	56.9	51.3	64.2	-
KPRTrack w/o prompt	61.5	58.4	65.8	3830
KPRTrack	**63.0** (+1.5)	**59.0** (+0.6)	**68.3** (+2.5)	**3344** (-486)

(b) Pose Tracking on PoseTrack21

Id	Main components of KPR						Dataset	
	PBH	MSF	Prompt ⊕	Prompt ⊖	BIPO	SOL	Occluded-PoseTrack R-1	Occluded-PoseTrack mAP
1							84.8	74.7
2	✓						83.3 (-1.5)	72.2 (-2.5)
3	✓	✓					84.8 (+0.0)	76.1 (+1.4)
4	✓	✓	✓				88.8 (+4.0)	80.4 (+5.7)
5	✓	✓	✓	✓			90.0 (+5.2)	80.7 (+6.0)
6	✓	✓			✓		86.1 (+1.3)	76.7 (+2.0)
7		✓	✓	✓	✓		80.4 (+5.1)	89.9 (+5.7)
8	✓	✓	✓	✓	✓		**92.3** (+6.5)	**82.3** (+7.6)
9	✓	✓	✓	✓	✓	✓	90.6 (+5.8)	81.2 (+6.5)

(c) Ablation study of KPR architecture

Qualitative Results



(a) Ranked gallery samples for a given query, when the input prompt is disabled/enabled. Green/red borders are correct/incorrect matches. The model focuses on the wrong target when there is no prompt.

(b) Part-attention maps for various input prompts.

(c) Pose tracking results: all track IDs, depicted by unique skeleton colors, are maintained after a long occlusion by the front cyclist.

Please check out our **Github** to get access to our codebase, trained models, dataset and annotations.

https://github.com/VlSomers/keypoint_promptable_reidentification

vladimir.somers@epfl.ch