Person Re-Identification

Yiheng Liu

Outline

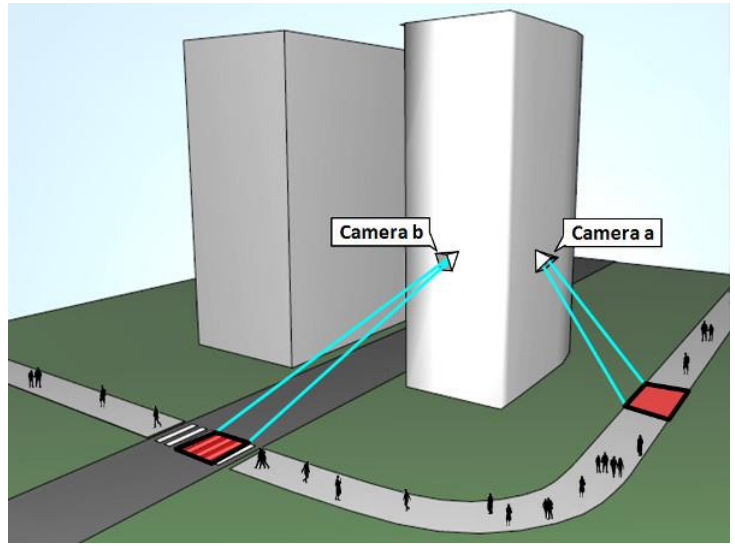
- Background
- Image-Based Person Re-Identification
- Partial Person Re-identification
- Video-Based Person Re-Identification
- Our Methods

Outline

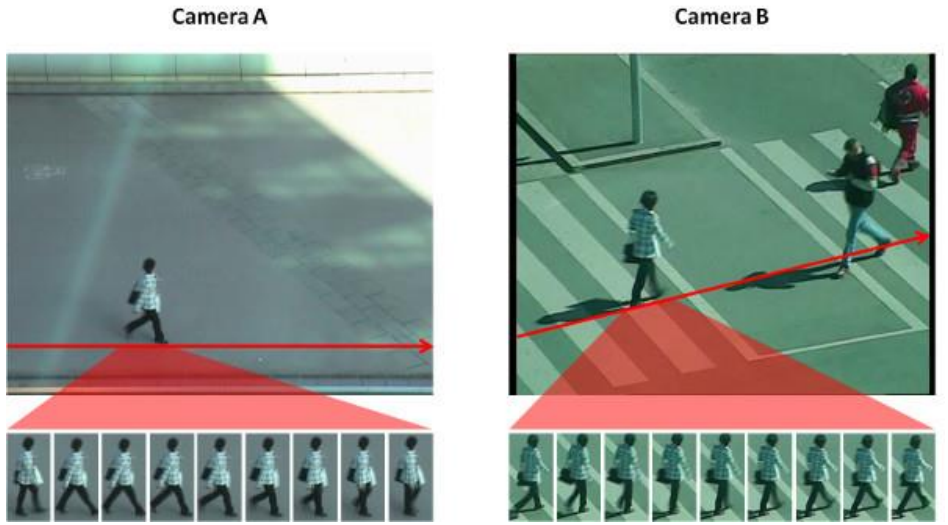
- Background
- Image-Based Person Re-Identification
- Partial Person Re-identification
- Video-Based Person Re-Identification
- Our Methods

Person Re-identification

- Person re-identification aims to match persons across non-overlapping surveillance camera views.



数据 检测 跟踪 匹配



Images：https://www.tugraz.at/institute/icg/research/team-bischof/lrs/downloads/prid11/

Datasets and Protocols

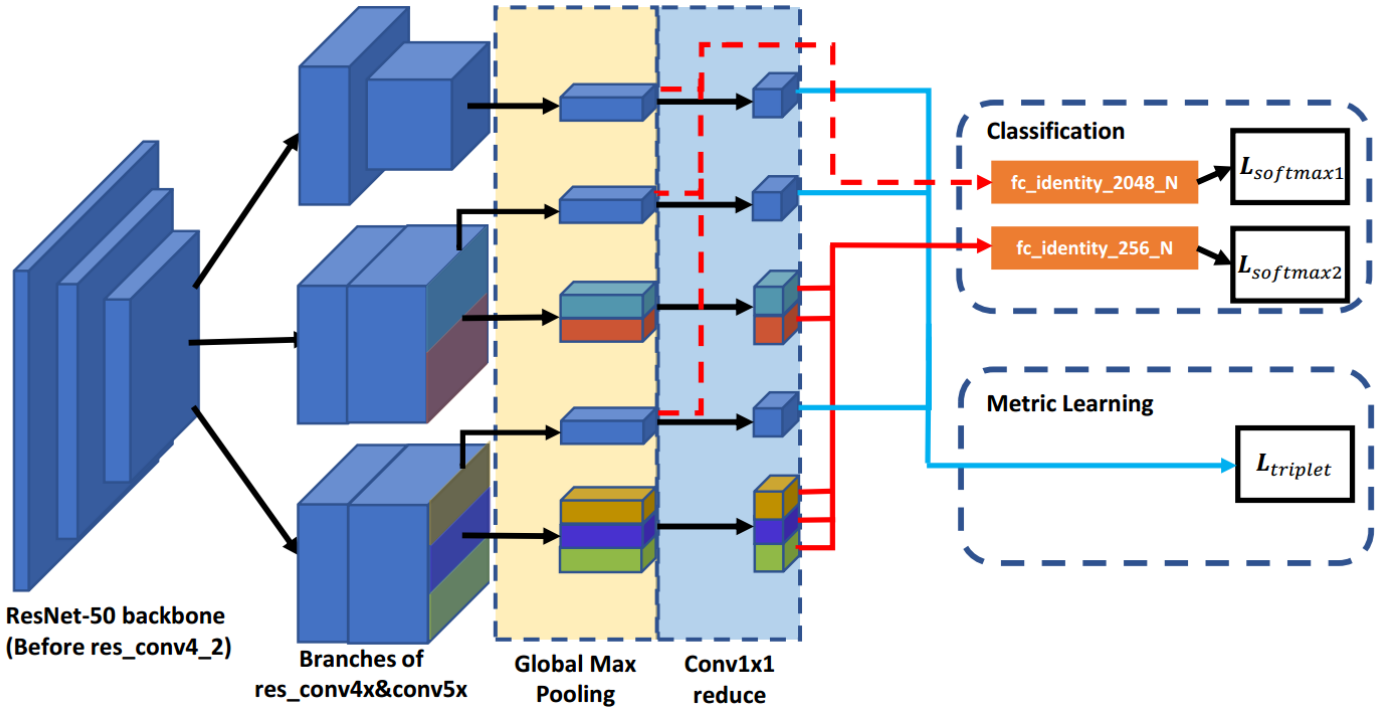
	Dataset	ID	Images
Image	Market1501	1501	32217
	DukeMTMC-reID	1812	36441
	CUHK03	1467	13164
	MSMT17	4101	126441
Video	iLIDS-VID	300	42495
	PRID2011	934	24541
	MARS	1261	1191003

- Cumulated Matching Characteristics (CMC)：Rank-1, Rank-5
- mean Average Precision (mAP)

Outline

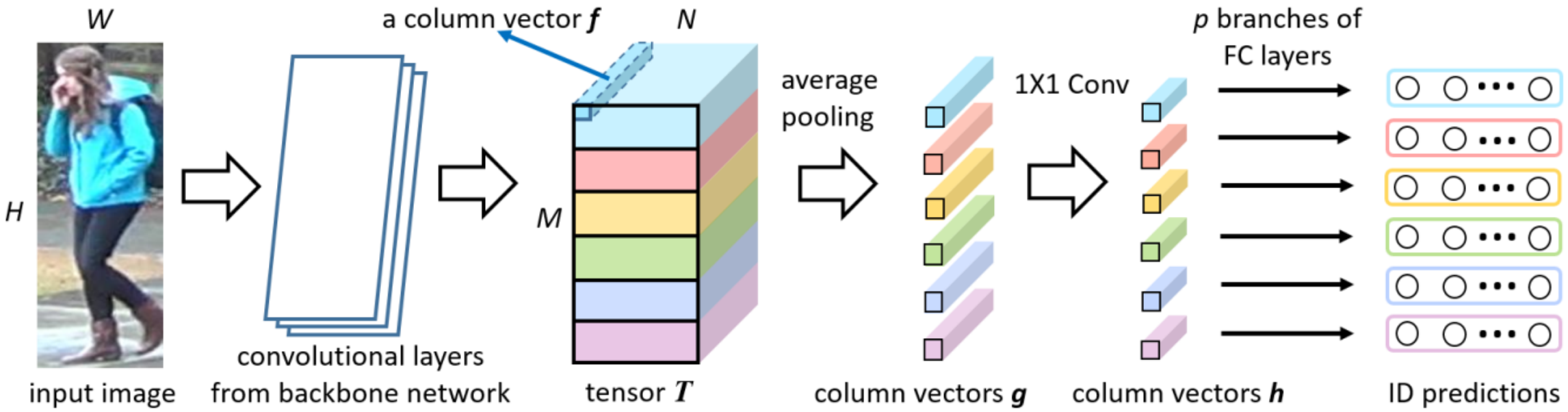
- Background
- Image-Based Person Re-Identification
- Partial Person Re-identification
- Video-Based Person Re-Identification
- Our Methods

Learning Discriminative Features with Multiple Granularities for Person Re-Identification

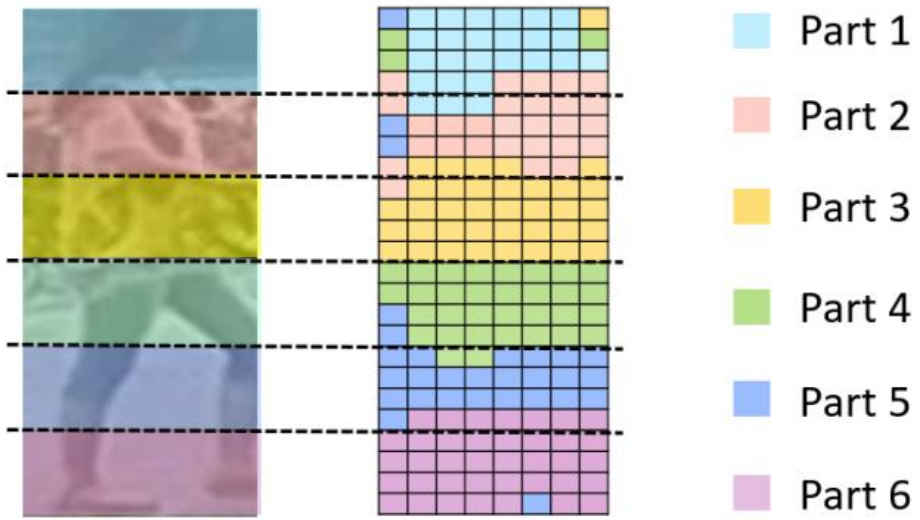


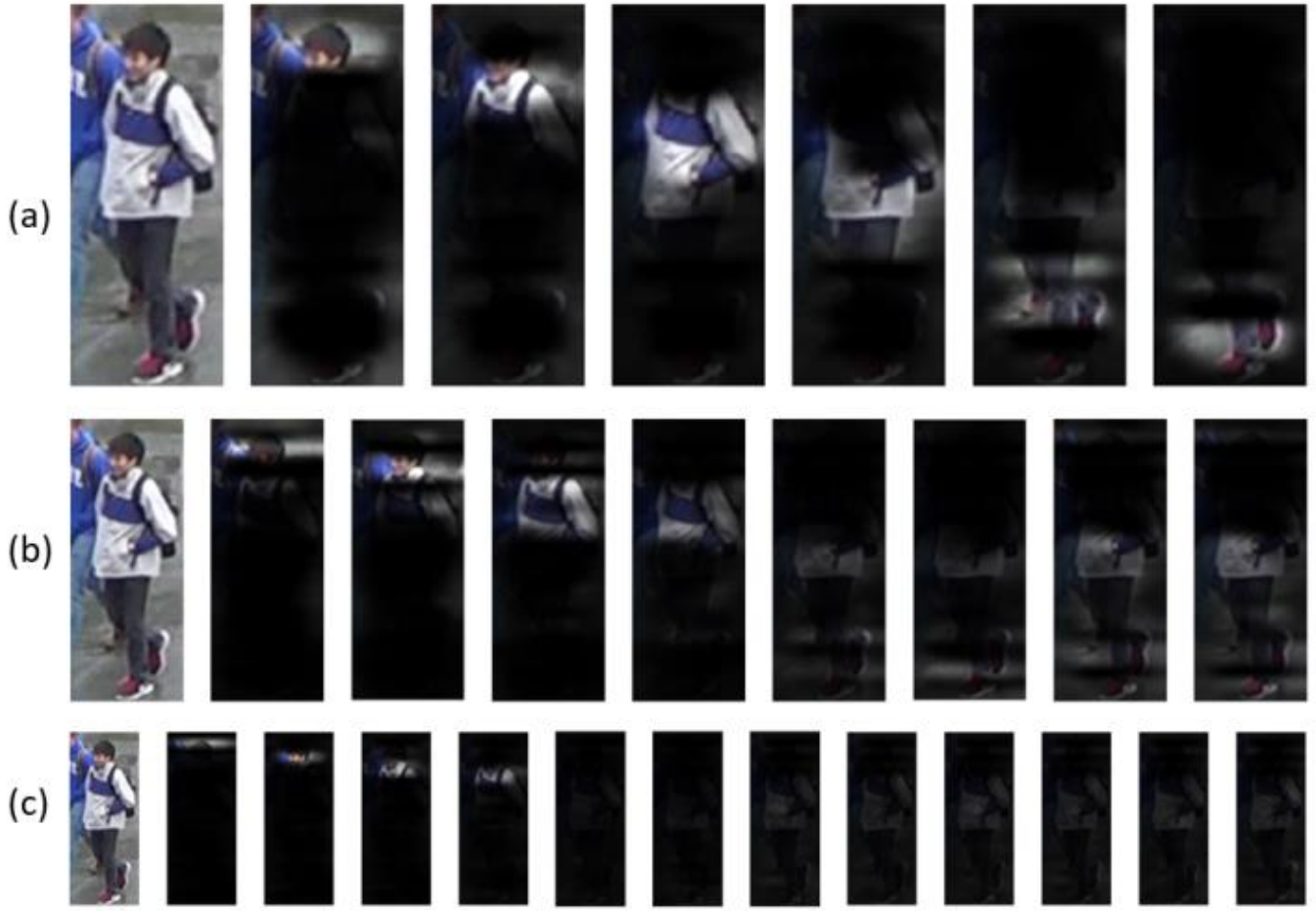
Model	Rank-1	Rank-5	Rank-10	mAP
ResNet-50	87.5	94.9	96.7	71.4
ResNet-101	90.4	95.7	97.2	78.0
ResNet-50+TP	88.7	96.0	97.2	75.0
Global (Branch)	89.8	95.8	97.5	78.5
Part-2 (Single)	92.6	97.1	98.0	80.2
Part-2 (Branch)	94.4	97.9	98.8	83.9
Part-3 (Single)	93.1	97.6	98.7	82.1
Part-3 (Branch)	94.4	98.2	98.8	84.1
G+P2+P3 (Single)	94.4	97.6	98.5	85.2
MGN w/o TP	95.3	97.9	98.7	86.2
MGN	**95.7**	**98.3**	**99.0**	**86.9**

Beyond Part Models: Person Retrieval with Refined Part Pooling (and A Strong Convolutional Baseline)



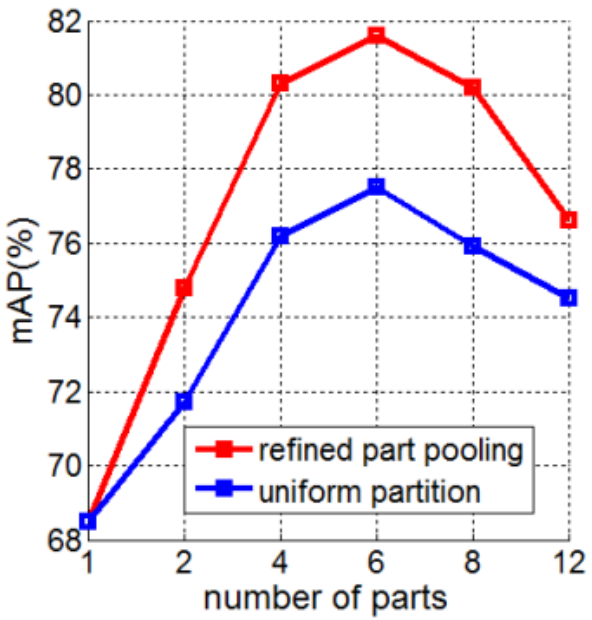
Refined Part Pooling

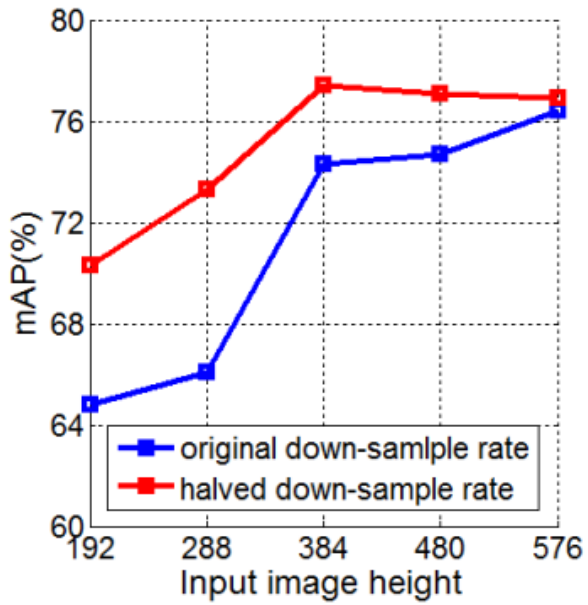


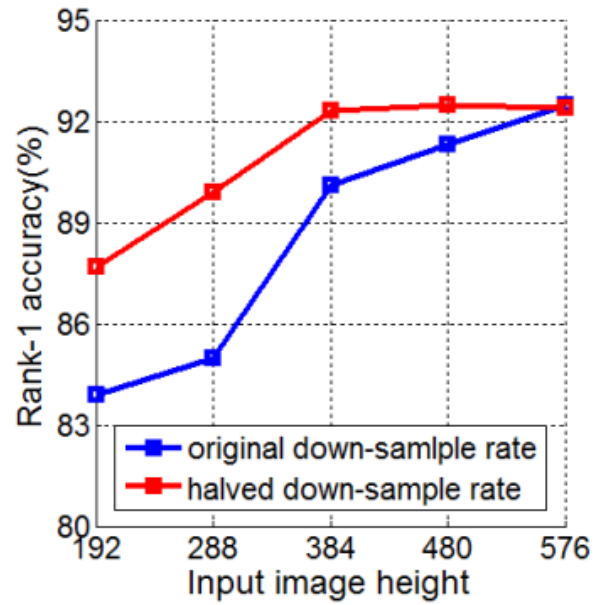


$$P_i = \{P(P_i|f) \times f, \forall f \in F\}$$

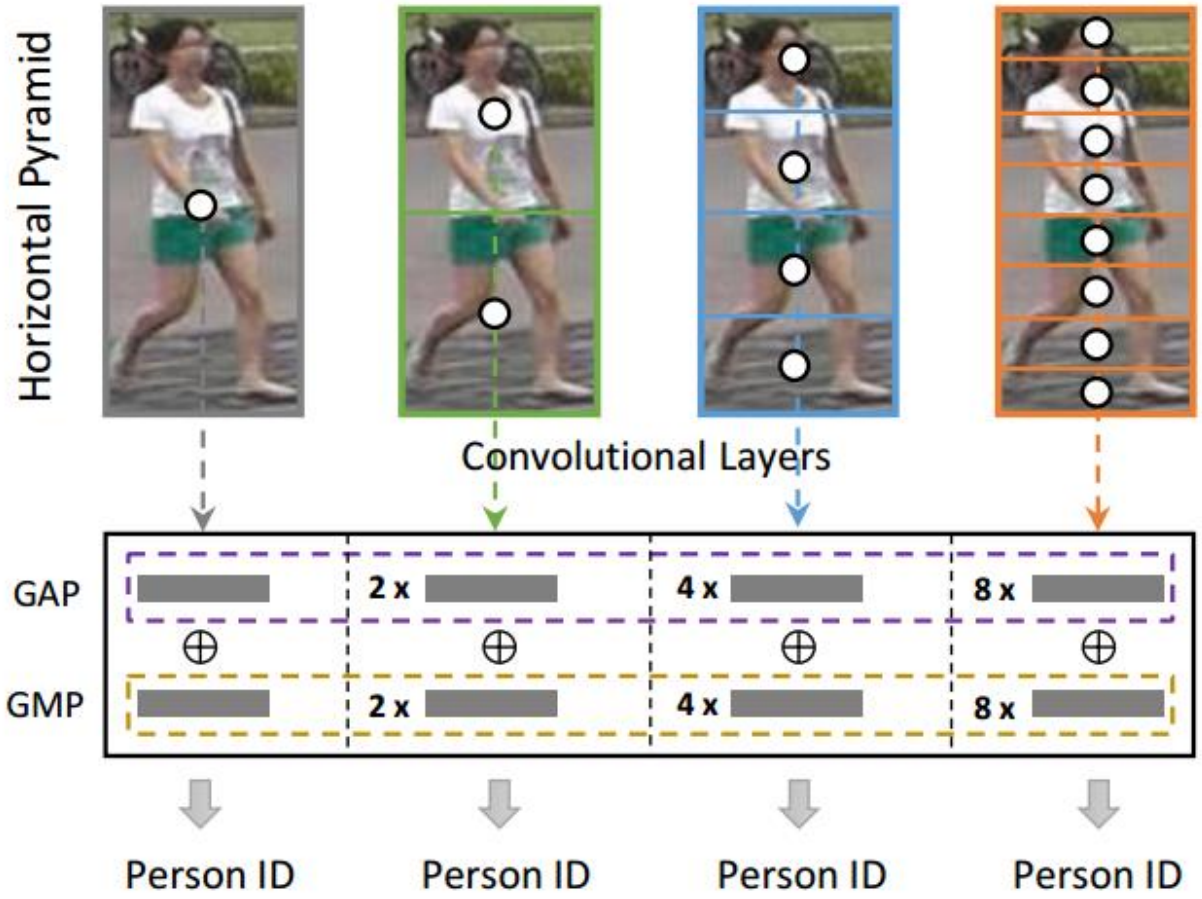
Models	Feature	dim	Market-1501				DukeMTMC-reID				CUHK03			
			R-1	R-5	R-10	mAP	R-1	R-5	R-10	mAP	R-1	R-5	R-10	mAP
IDE	pool5	2048	85.3	94.0	96.3	68.5	73.2	84.0	87.6	52.8	43.8	62.7	71.2	38.9
IDE	FC	256	83.8	93.1	95.8	67.7	72.4	83.0	87.1	51.6	43.3	62.5	71.0	38.3
Variant 1	$\mathcal{G}$	12288	86.7	95.2	96.5	69.4	73.9	84.6	88.1	53.2	43.6	62.9	71.3	38.8
Variant 1	$\mathcal{H}$	1536	85.6	94.3	96.3	68.3	72.8	83.3	87.2	52.5	44.1	63.0	71.5	39.1
Variant 2	$\mathcal{G}$	12288	91.2	96.6	97.7	75.0	80.2	88.8	91.3	62.8	52.6	72.4	80.9	45.8
Variant 2	$\mathcal{H}$	1536	91.0	96.6	97.6	75.3	80.0	88.1	90.4	62.6	54.0	73.7	81.4	47.2
PCB	$\mathcal{G}$	12288	92.3	97.2	98.2	77.4	81.7	89.7	91.9	66.1	59.7	77.7	85.2	53.2
PCB	$\mathcal{H}$	1536	92.4	97.0	97.9	77.3	81.9	89.4	91.6	65.3	61.3	78.6	85.6	54.2
PCB+RPP	$\mathcal{G}$	12288	**93.8**	**97.5**	**98.5**	**81.6**	**83.3**	**90.5**	**92.5**	**69.2**	62.8	79.8	86.8	56.7
PCB+RPP	$\mathcal{H}$	1536	93.1	97.4	98.3	81.0	82.9	90.1	92.3	68.5	**63.7**	**80.6**	**86.9**	**57.5**

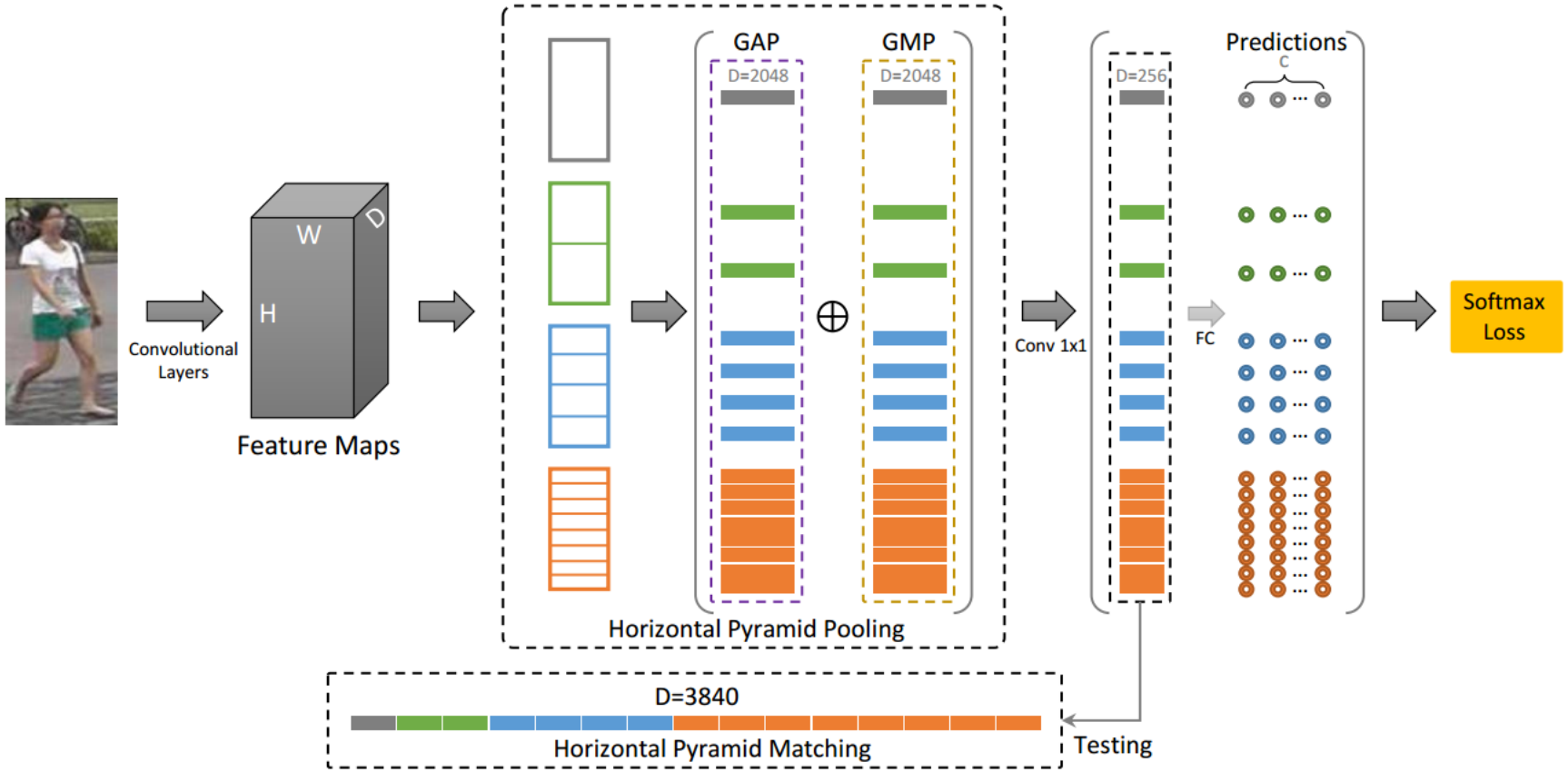






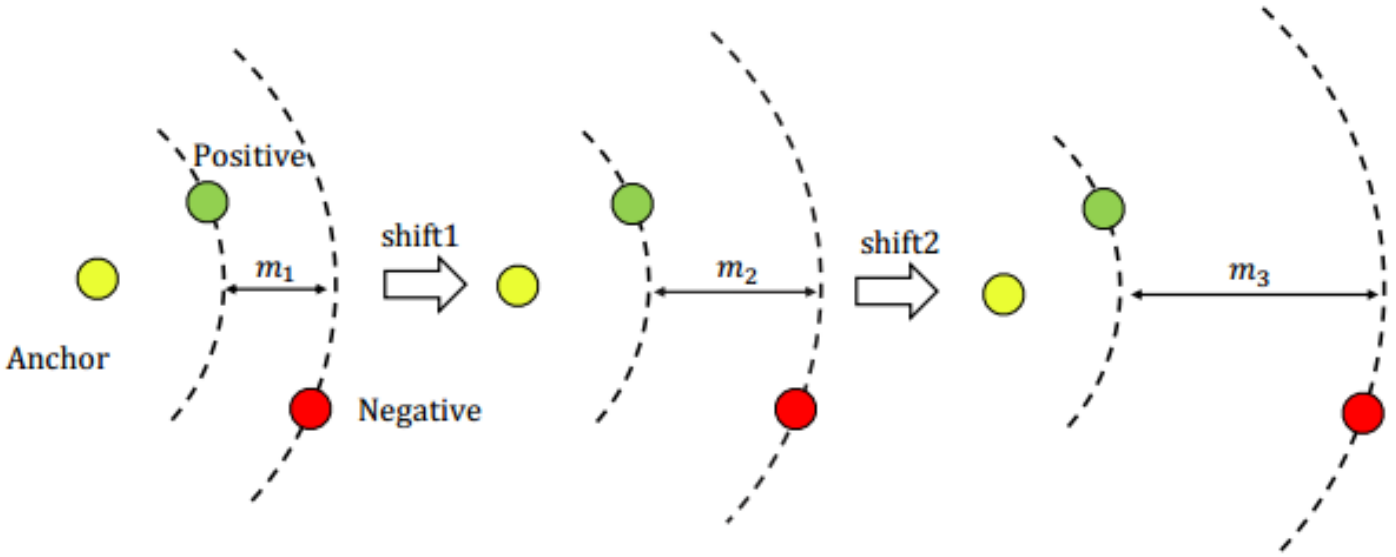
Horizontal Pyramid Matching for Person Re-identification

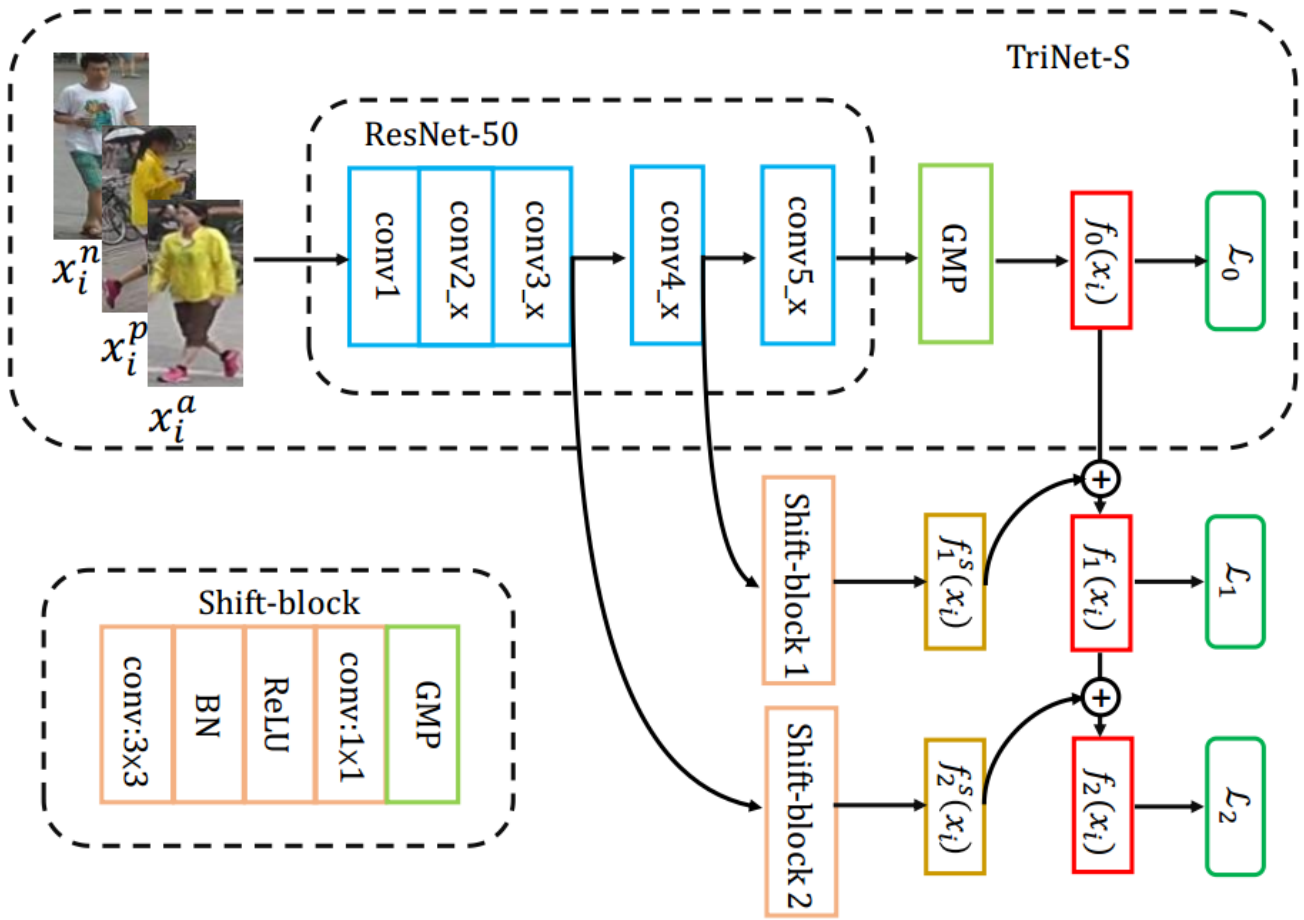




Model	Feature Dim	Market1501				DukeMTMC-ReID				CUHK03			
		R1	R5	R10	mAP	R1	R5	R10	mAP	R1	R5	R10	mAP
HPM + #PS 1 + Avg pool	256	88.1	94.6	96.4	71.2	79.3	89.7	91.9	61.0	39.2	61.1	71.6	37.3
HPM + #PS 2 + Avg pool	256x(1+2)	92.0	96.9	97.9	78.3	83.1	91.9	93.4	68.9	53.2	73.2	79.6	48.9
HPM + #PS 3 + Avg pool	256x(1+2+4)	92.3	97.2	97.9	79.3	84.5	92.4	94.1	70.8	58.2	76.7	83.1	52.8
HPM + #PS 4 + Avg pool	256x(1+2+4+8)	93.2	97.3	98.1	79.5	84.8	92.5	94.1	72.1	58.6	76.8	83.8	53.4
HPM + #PS 4 + Max pool	256x(1+2+4+8)	93.6	**97.7**	98.3	81.6	86.2	**93.2**	94.8	74.1	62.4	78.9	**86.3**	57.4
HPM + #PS 4 + Max+Avg pool	256x(1+2+4+8)	**94.2**	97.5	**98.5**	**82.7**	**86.6**	93.0	**95.1**	**74.3**	**63.9**	**79.7**	86.1	**57.5**

Learning Incremental Triplet Margin for Person Re-identification





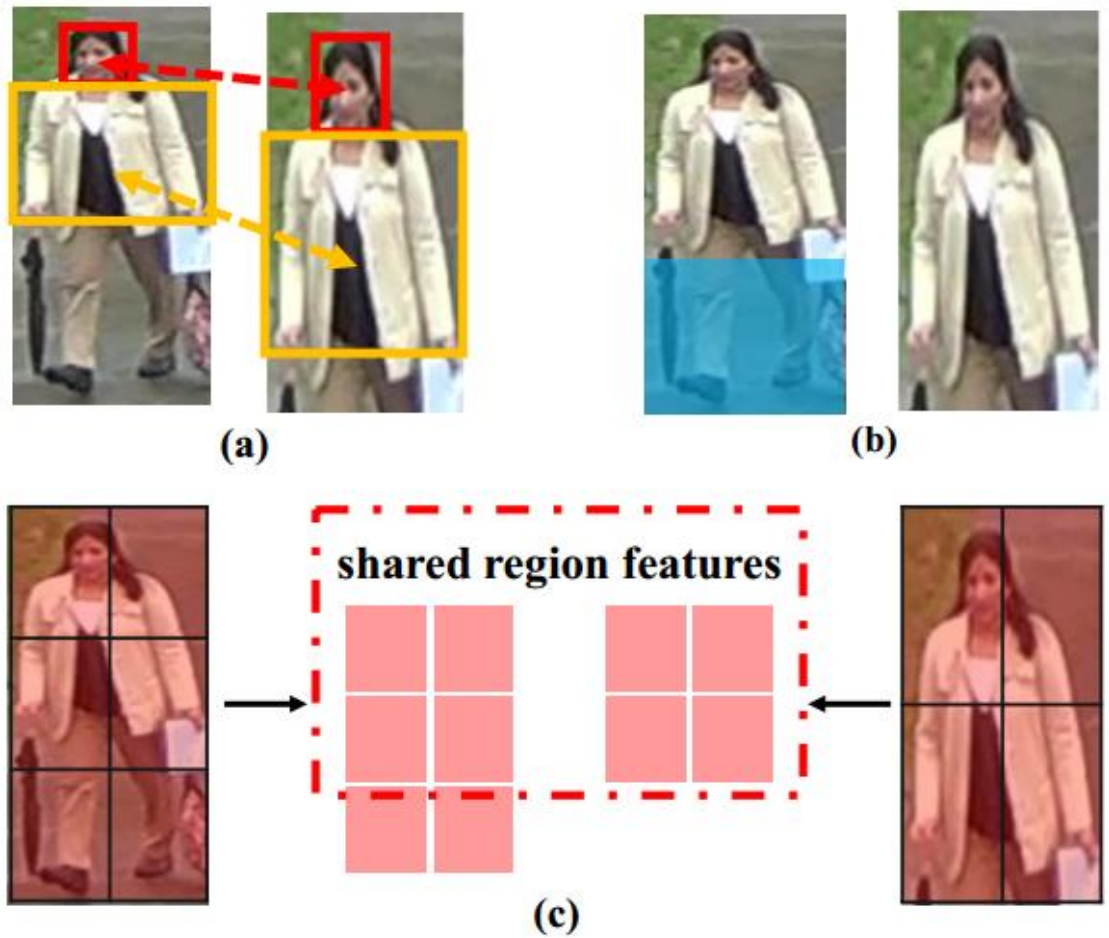
	m	Rank-1	Rank-5	Rank-10	mAP
TriNet-S	1	90.1	94.8	96.5	77.9
	4	90.9	96.6	97.5	79.1
	7	90.2	96.4	97.0	78.2
	10	89.9	95.2	96.2	77.7
$f_0(\cdot)$	4	92.1	96.9	98.0	80.9
$f_1(\cdot)$	7	92.6	97.1	98.5	82.2
$f_2(\cdot)$	10	92.6	97.5	98.5	82.3

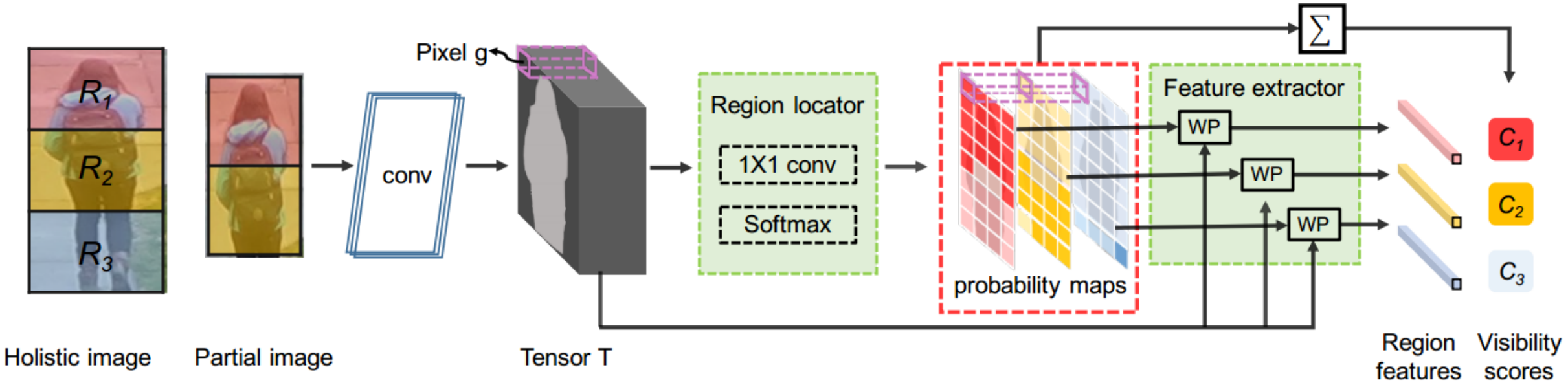
Measure (%)	Rank-1	Rank-5	Rank-10	mAP
LITM-C5C5C5	92.0	97.0	98.2	81.2
LITM-C3C4C5	90.8	96.3	97.9	79.4
LITM-C5C4C3	92.6	97.5	98.5	82.3

Outline

- Background
- Image-Based Person Re-Identification
- Partial Person Re-identification
- Video-Based Person Re-Identification
- Our Methods

Perceive Where to Focus: Learning Visibility-aware Part-level Features for Partial Person Re-identification





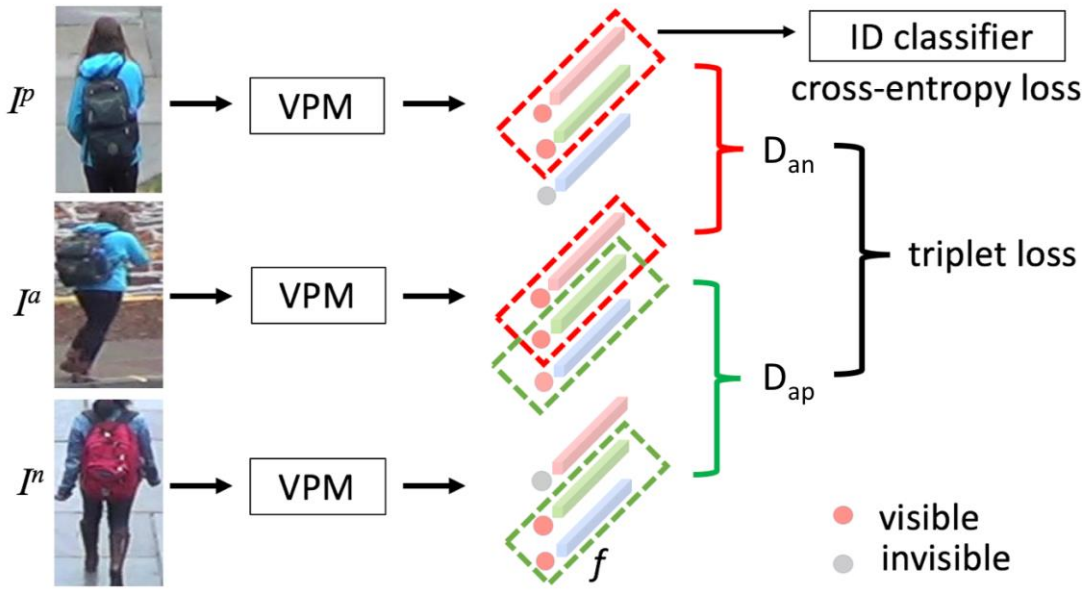
Region feature:

$$f_i = \frac{\sum_{g \in T} P(R_i|g)g}{C_i}, \forall i \in \{1, 2, \cdots, p\}$$

Overall distance:

$$D^{kl} = \frac{\sum_{i=1}^{p} C_i^k C_i^l D_i^{kl}}{\sum_{i=1}^{p} C_i^k C_i^l}$$

$$D_i^{kl} = \|f_i^k - f_i^l\|$$



- The training of region locator

$$L_R = -\sum_{g \in T} \mathbb{1}_{i=L} log(P(R_i|g))$$

- Cross-entropy loss

$$L_{ID} = -\sum_{i \in V} \mathbb{1}_{k=y} log(softmax(IP_i(f_i)))$$

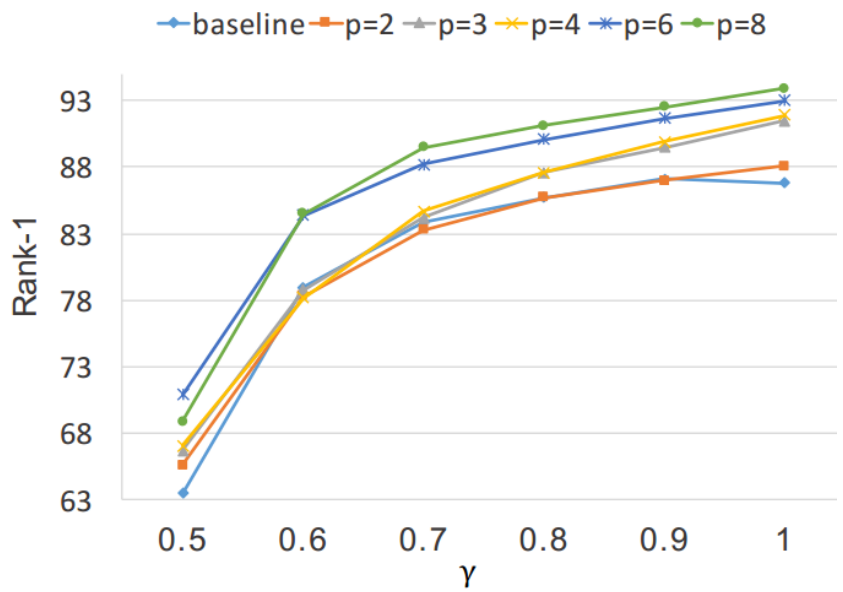
- Triplet loss

$$L_{tri} = [D^{ap} - D^{an} + \alpha]_+ ,$$

$$D^{ap} = \frac{\sum_{i \in (V^a \cap V^p)} \|f_i^a - f_i^p\|}{|V^a \cap V^p|},$$

$$D^{an} = \frac{\sum_{i \in (V^a \cap V^n)} \|f_i^a - f_i^n\|}{|V^a \cap V^n|},$$

Dataset	γ	baseline				PCB				VPM			
		R-1	R-5	R-10	mAP	R-1	R-5	R-10	mAP	R-1	R-5	R-10	mAP
Market-1501	0.5	64.5	82.2	88.1	44.4	0.9	3.2	5.6	1.7	70.9	86.5	92.1	48.8
	0.6	79.0	91.4	94.3	57.9	8.1	16.5	23.2	6.6	84.4	94.3	96.1	62.5
	0.7	83.9	93.9	95.9	63.7	36.5	58.9	67.4	26.8	88.2	95.8	97.2	71.7
	0.8	85.7	94.3	96.4	66.1	71.9	87.3	91.4	56.8	90.1	95.8	97.7	74.7
	0.9	87.1	95.5	97.4	67.8	88.8	95.8	97.1	77.2	91.7	96.6	98.0	78.7
	1.0	86.8	95.3	97.4	67.7	93.4	97.8	98.4	83.0	93.0	97.8	98.8	80.8
DukeMTMC-reID	0.5	65.0	81.1	86.7	47.2	5.0	10.1	13.6	4.0	69.5	83.1	87.9	52.2
	0.6	76.2	87.3	90.4	55.4	13.1	25.6	33.5	10.5	78.2	89.0	91.3	60.9
	0.7	76.3	87.3	90.6	90.6	35.9	57.0	65.4	28.4	80.3	89.5	92.0	63.1
	0.8	76.3	88.3	91.9	58.8	64.0	82.6	87.7	52.3	80.3	89.3	92.4	63.5
	0.9	77.0	88.1	91.7	59.0	81.6	90.4	93.0	70.3	81.7	90.9	93.1	70.7
	1.0	76.2	87.3	91.2	58.6	84.1	92.4	94.5	73.2	83.6	91.7	94.2	72.6



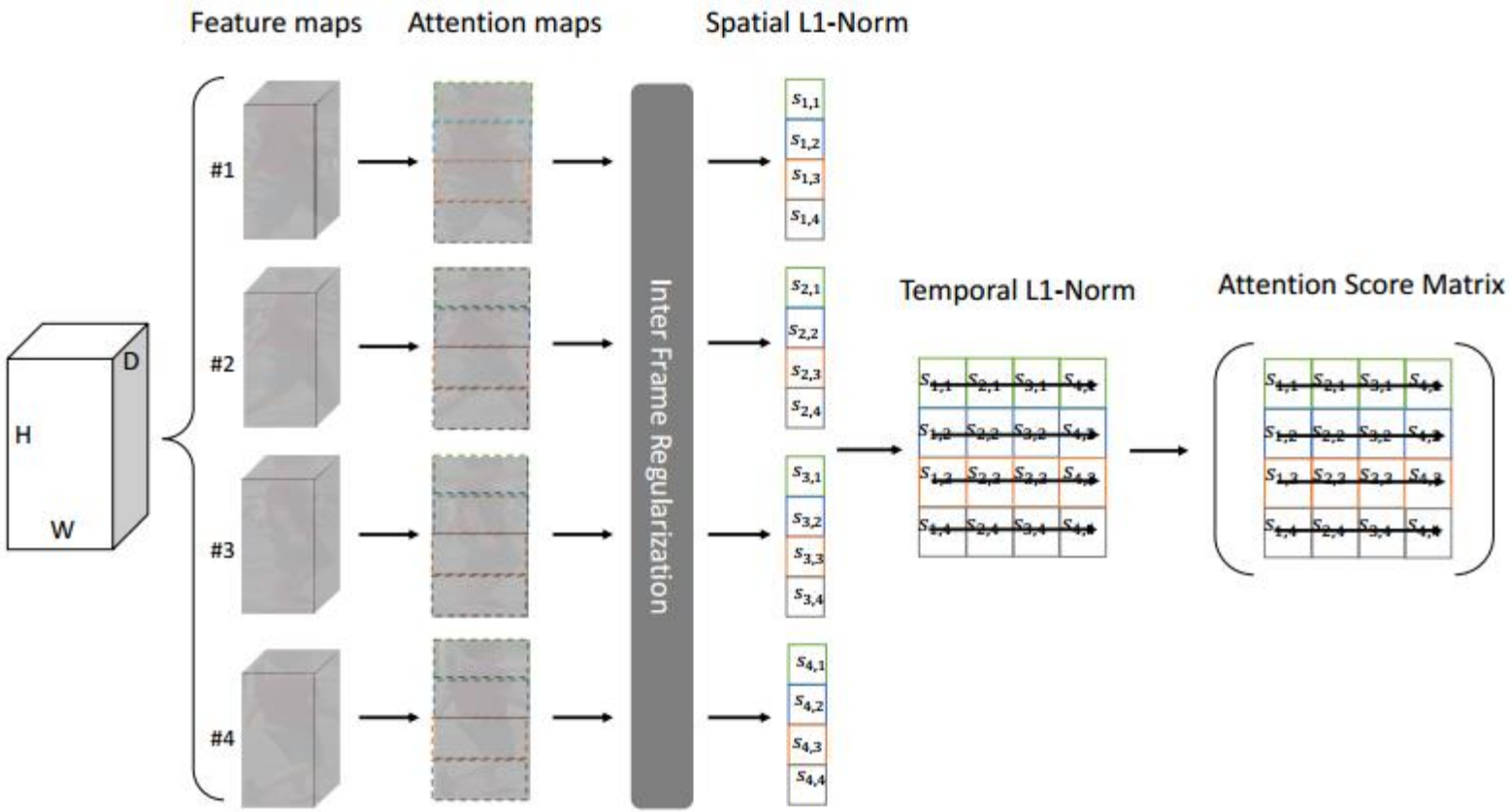
Methods	Partial-REID		Partial-iLIDS	
	R-1	R-3	R-1	R-3
MTRC [15]	23.7	27.3	17.7	26.1
AMC+SWM [36]	37.3	46.0	21.0	32.8
DSR [7]	50.7	70.0	58.8	67.2
SFR [8]	56.9	78.5	63.9	74.8
VPM (Bottom)	53.2	73.2	53.6	62.3
VPM (Top)	64.3	**83.6**	**67.2**	**76.5**
VPM (Bilateral)	**67.7**	81.9	65.5	74.8

Methods	Partial-iLIDS			Market-1501		
	R-1	R-3	R-5	R-1	R-5	mAP
VPM.	67.2	76.5	82.4	93.0	97.8	80.8
VPM (no triplet)	57.1	73.9	79.0	91.3	97.0	77.8
MVPM-1	63.0	74.8	82.4	93.0	96.3	79.7
MVPM-2	61.3	73.1	79.0	92.8	97.4	80.1
MVPM-3	58.8	74.8	82.4	91.4	96.5	75.5
MVPM-4	59.7	74.8	78.2	90.4	96.6	75.7

Outline

- Background
- Image-Based Person Re-Identification
- Partial Person Re-identification
- Video-Based Person Re-Identification
- Our Methods

STA: Spatial-Temporal Attention for Large-Scale Video-based Person Re-Identification

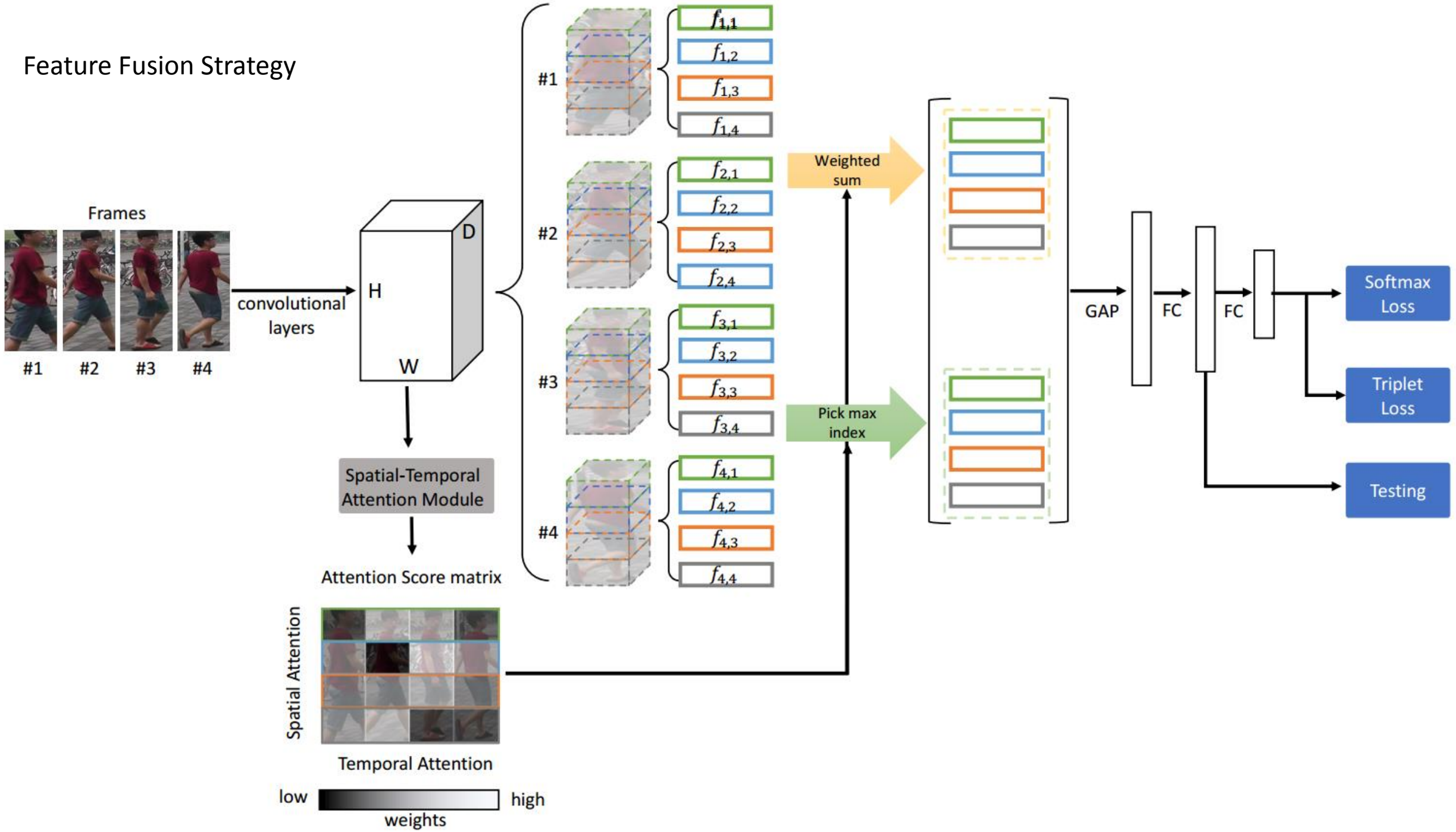


Inter-Frame Regularization

$$Reg = ||g_i - g_j||_F$$

$$= \sqrt{\sum_{h=1}^{H}\sum_{w=1}^{W} |g_i(h,w) - g_j(h,w)|^2}$$

Feature Fusion Strategy



Model	MARS				DukeMTMC-VideoReID			
	R1	R5	R10	mAP	R1	R5	R10	mAP
Baseline	74.5	88.8	91.8	64.0	79.1	93.9	96.0	76.8
Baseline + TL	80.8	92.1	94.3	74.0	90.6	95.8	96.7	89.7
Baseline + TL + Avg	82.5	92.9	94.9	75.0	91.8	97.4	98.0	91.0
Baseline + TL + STA	84.8	94.6	96.2	78.0	93.3	98.1	98.6	92.7
Baseline + TL + STA + Fusion	85.3	95.1	96.4	79.1	95.3	98.1	99.1	93.9
Baseline + TL + STA + Fusion + Reg	86.3	95.7	97.1	80.8	96.2	99.3	99.6	94.9

Sequence Length	MARS				DukeMTMC-VideoReID			
	R1	R5	R10	mAP	R1	R5	R10	mAP
N=2	81.7	93.8	95.7	75.7	90.3	97.6	98.6	89.0
N=4	86.3	95.7	97.1	80.8	96.2	99.3	99.6	94.9
N=6	86.2	95.7	96.9	81.0	96.0	99.4	99.7	95.0
N=8	86.2	95.7	97.1	81.2	96.0	99.3	99.6	95.0

Number of Spatial Regions	MARS				DukeMTMC-VideoReID			
	R1	R5	R10	mAP	R1	R5	R10	mAP
K=2	85.3	95.1	96.6	80.3	94.7	99.0	99.6	93.8
K=4	86.3	95.7	97.1	80.8	96.2	99.3	99.6	94.9
K=8	85.5	95.3	96.9	80.4	95.2	99.1	99.4	93.8

THANK YOU FOR LISTENING