Latent Distribution Alignment for Domain Generalizable Person Re-identification

Ren Nie[1], Jin Ding[1], Lingxiao He[2], Xue Zhou[1,3] *

[1] School of Automation Engineering, University of Electronic Science and Technology of China

[2] Meituan, China

[3] Shenzhen Institute for Advanced Study, UESTC

Motivation

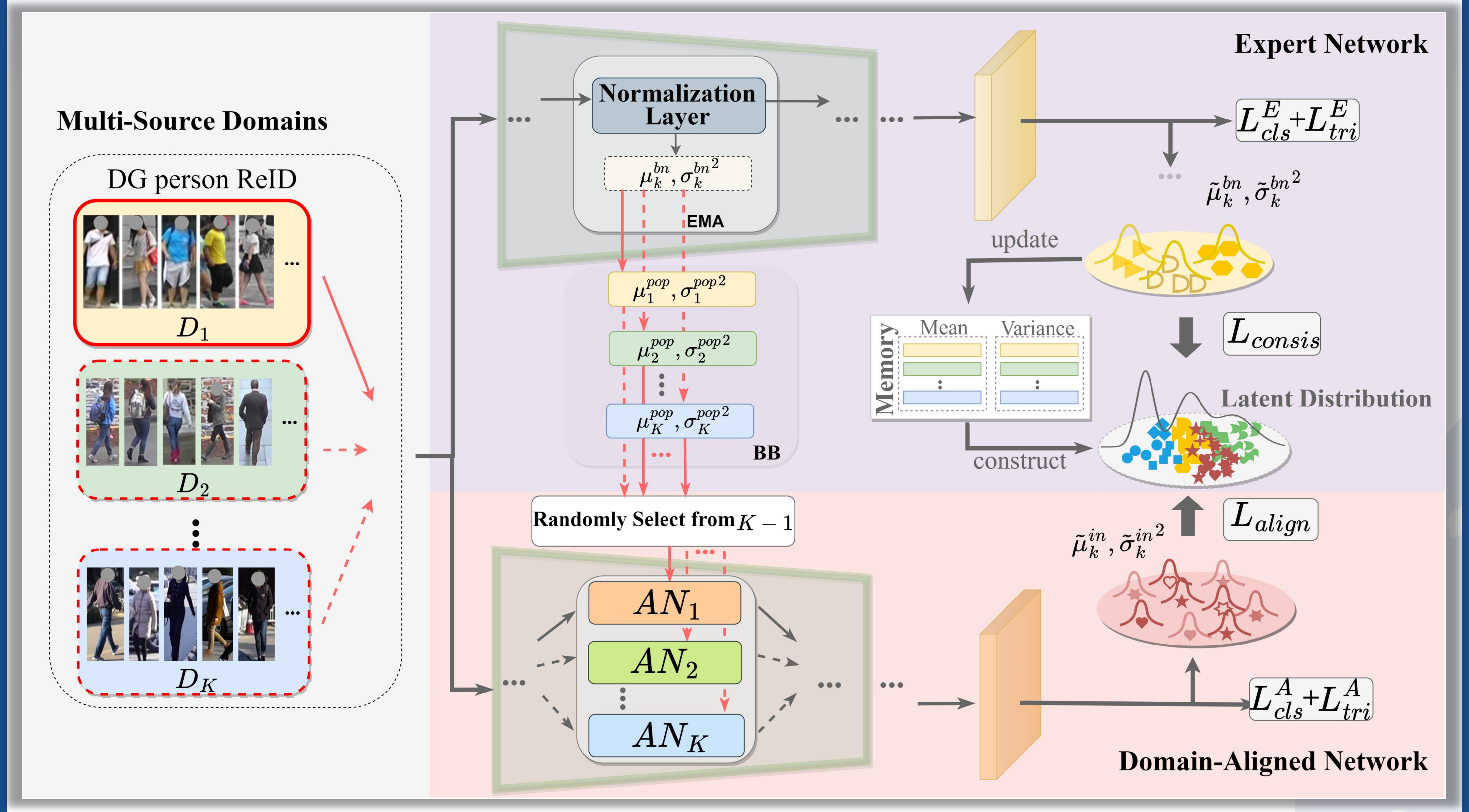
Domain Generalizable Person Re-identification (DG ReID) aims to solve the domain shift problem between training source domains and unseen target domains.

- **Previous methods endeavored to extract domain-invariant information from source domains without directly accounting for the absence of target domains.**
- **The BatchNorm (BN) statistics of feature maps contain a wealth of domain-specific representation information.**
- **By leveraging the domain-specific information, we can regard each source domain as a mimical target domain of the remaining source domains.**

Contribution

- ➢ **We propose a novel Latent Distribution Alignment (LDA) framework for DG ReID which dynamically construct a latent distribution to align BN statistics of source domains and IN statistics of target domains.**
- ➢ **We introduce the Buffer Bank (BB) and Anti-Normalization (AN), which treat each source domain as the mimical target domain of remaining source domains without any added learnable parameters.**
- ➢ **Extensive experiments demonstrate the simplicity yet effectiveness and comparable or even better state-of-the-art performance of our LDA framework.**

Methodology



Overview of our proposed LDA framework

- ➢ **We construct an latent distribution by using the BN statistics of each source domain from the output of the expert network backbone.Then, we introduce a consistency loss (L_{consis}) to regulate the distribution of the output BN statistics from the Expert Network(-E) and an alignment loss (L_{align}) to align the output IN statistics from Domain-Aligned Network(-A) with the latent distribution.**
- ➢ **BB is used within Network-E to store the domain-specific BN information for each source domain across every normalization layer.**
- ➢ **With the help of BB, AN treats the input source domain of Network-A as an unseen target domain and transfer the distribution to another chosen source domain as follows:**

$$AN_k(X_k[h,w]) = \frac{X_k[h,w] - \mu_k^{in}}{\sqrt{\sigma_k^{in^2} + \epsilon}} \left(\sqrt{\sigma_j^{pop^2} + \epsilon}\right) + \mu_j^{\text{pop}}, j \sim \{1, 2, ..., K\} \backslash \{k\}$$

Experiments

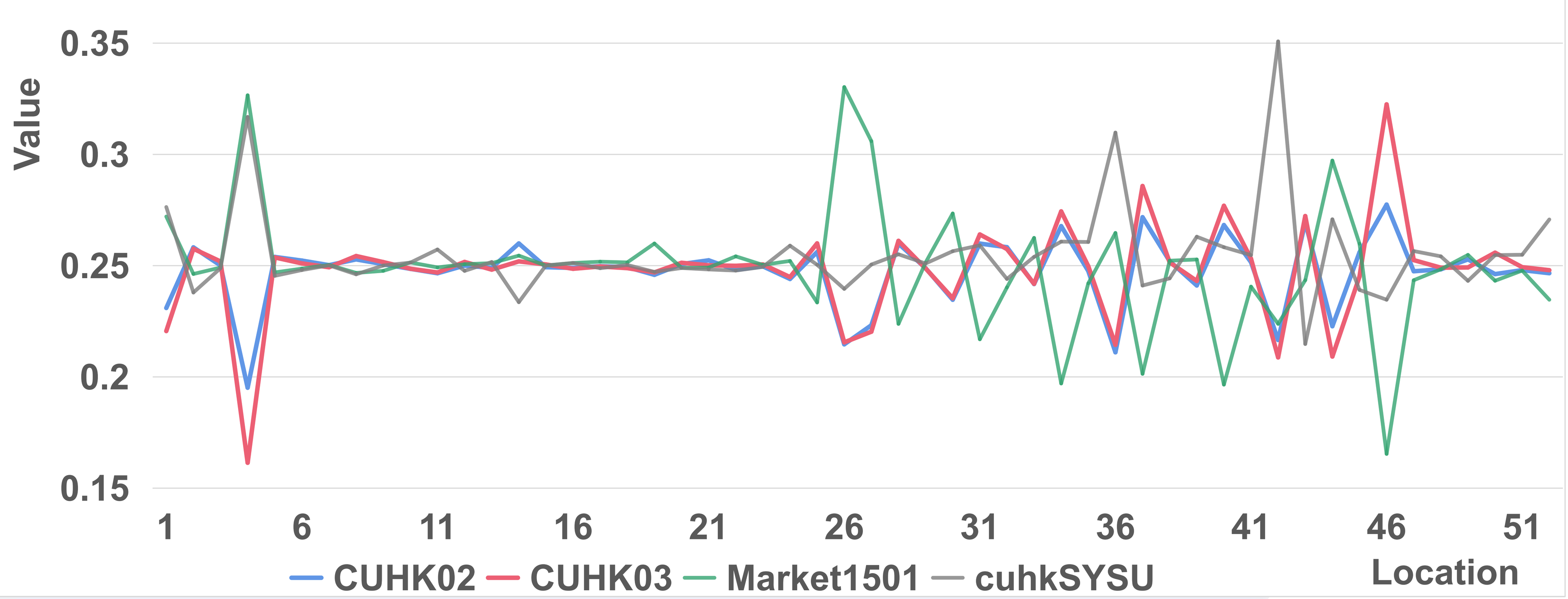
Comparision with the SOTAS under Protocol-1

Method	Reference	Source domain	Target domain: PRID mAP	PRID Rank-1	GRID mAP	GRID Rank-1	VIPeR mAP	VIPeR Rank-1	iLIDs mAP	iLIDs Rank-1	Average mAP	Average Rank-1
SNR [9]	CVPR'20	M+D+C2 +C3+CS	66.5	52.1	47.7	40.2	61.3	52.9	89.9	84.1	66.4	57.3
RaMoE [8]	CVPR'21		67.3	57.7	54.2	46.8	64.6	56.6	**90.2**	**85.0**	62.0	61.5
MDA [12]	CVPR'22		–	–	62.9	**61.2**	71.7	63.5	84.4	80.4	–	–
* $QAConv_{50}$ [17]	ECCV'20	M+C2 +C3+CS	62.2	52.3	57.4	48.6	66.3	57.0	81.9	75.0	67.0	58.2
* M^3L [6]	CVPR'21		65.3	55.0	50.5	40.0	68.2	60.8	74.3	65.0	64.6	55.2
* MetaBIN [7]	CVPR'21		70.8	61.2	57.9	50.2	64.3	55.9	82.7	74.7	68.9	60.5
META [10]	ECCV'22		71.7	61.9	60.1	52.4	68.4	61.5	83.5	79.2	70.9	63.8
ACL [13]	ECCV'22		73.4	63.0	**65.7**	55.2	**75.1**	**66.4**	86.5	81.8	**75.2**	66.6
LDA (Ours)	This paper		**76.9**	**69.5**	63.3	53.8	73.7	65.2	84.3	79.0	74.7	**67.1**

Comparision with the SOTAS under Protocol-2&3

Method	Reference	Setting	M+MS+CS→ C3 mAP	M+MS+CS→ C3 Rank-1	M+CS+C3 → MS mAP	M+CS+C3 → MS Rank-1	MS+CS+C3 → M mAP	MS+CS+C3 → M Rank-1	Average mAP	Average Rank-1
* SNR [9]	CVPR'20	Protocol-2	8.9	8.9	6.8	19.9	34.6	62.7	16.8	30.5
* $QAConv_{50}$ [17]	ECCV'20		25.4	24.8	16.4	45.3	63.1	83.7	35.0	51.3
* M^3L [6]	CVPR'21		20.9	31.9	15.9	36.9	58.4	79.9	31.7	49.6
* MetaBIN [7]	CVPR'21		28.8	28.1	17.8	40.2	57.9	80.1	34.8	49.5
META [10]	ECCV'22		36.3	35.1	22.5	49.9	67.5	86.1	42.1	57.0
ACL [13]	ECCV'22		41.2	41.8	20.4	45.9	**74.3**	**89.3**	45.3	59.0
LDA (Ours)	This paper		**42.5**	**43.3**	**23.9**	**51.1**	70.0	86.9	**45.5**	**60.4**
* SNR [9]	CVPR'20	Protocol-3	17.5	17.1	7.7	22.0	52.4	77.8	25.9	39.0
* $QAConv_{50}$ [17]	ECCV'20		32.9	33.3	17.6	46.6	66.5	85.0	39.0	55.0
* M^3L [6]	CVPR'21		32.3	33.8	16.2	36.9	61.2	81.2	36.6	50.6
* MetaBIN [7]	CVPR'21		43.0	43.1	18.8	41.2	67.2	84.5	43.0	56.3
META [10]	ECCV'22		47.1	46.2	24.4	52.1	76.5	90.5	49.3	62.9
ACL [13]	ECCV'22		**49.4**	**50.1**	21.7	47.3	**76.8**	**90.6**	49.3	62.7
LDA (Ours)	This paper		48.4	49.4	**25.9**	**53.6**	74.8	89.6	**49.7**	**64.2**

Visualization of the Buffer Bank under Protocol-1



It is evident that the curves for CUHK02 and CUHK03, which are from the same scene, are highly similar. The differences among other datasets are quite apparent. This observation indicates that our BB retains rich domain specific representation information.