Style Normalization and Restitution for Generalizable Person Re-identification

(CVPR 2020)

Xin et al.

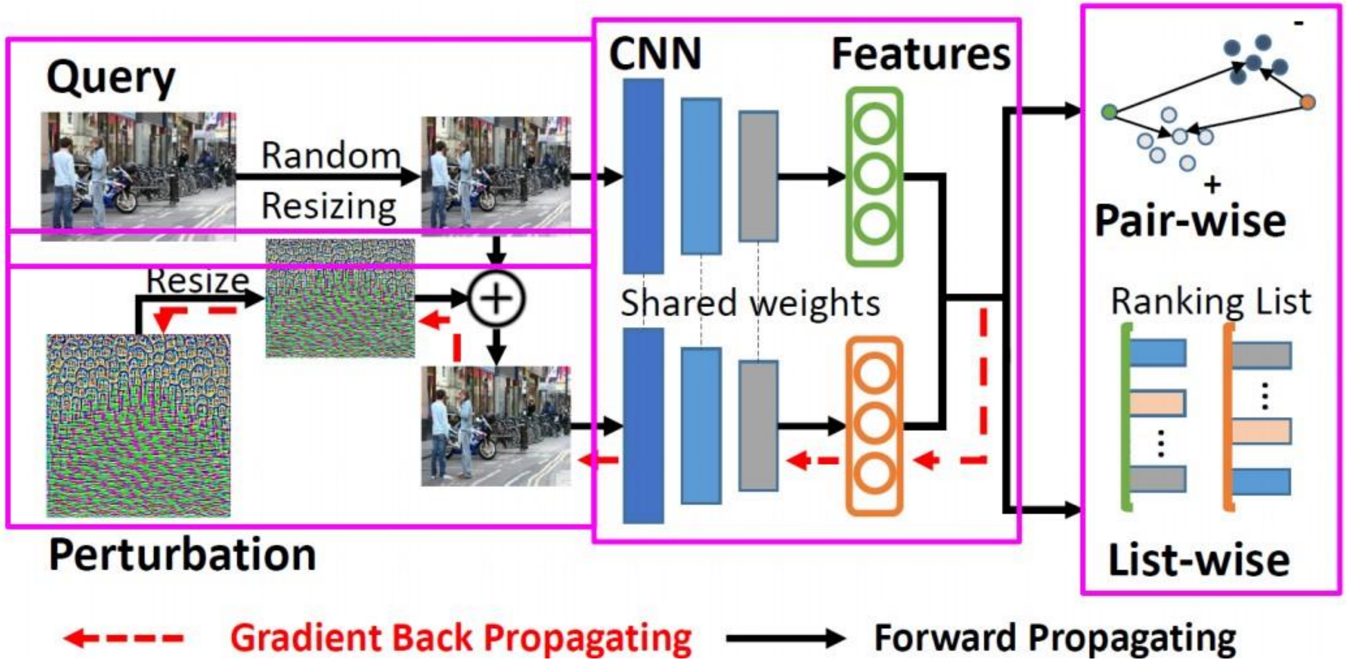
Presenter : Dongjun Kim

Contents

Review

Universal Perturbation Attack against Image Retrieval

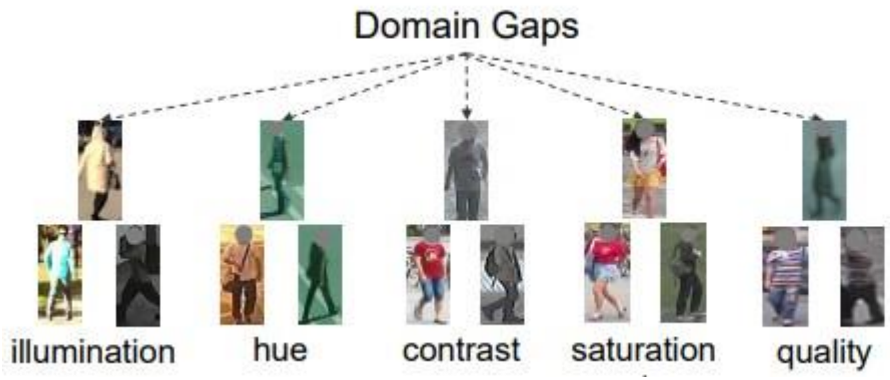
- **Aim :** Build UAP(Universal Adversarial Perturbations) specific to image retrieval task
- Mainly used **List-wise Loss** and **Pair-wise Loss** to perturb Triplet loss and entire ranking list



Introduction & Motivation

Introduction & Motivation

- Fully-supervised person re-identification (ReID) methods usually suffer from **poor generalization capability**
- The images are captured by different cameras under different environments
- Presents **large discrepancy** in terms of illumination, hue, contrast, saturation, quality



Introduction & Motivation

Introduction & Motivation

For generalizable ReID framework which achieves,

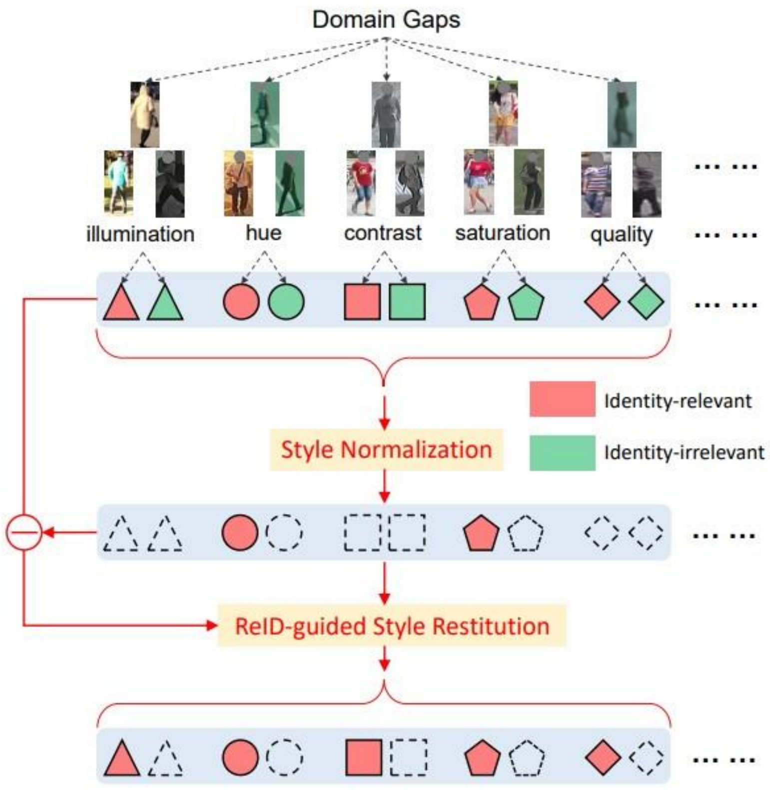
- Generalization capability
- Discrimination capability

→ **Disentangle** the identity-relevant features and identity-irrelevant features

Introduction & Motivation

Introduction & Motivation

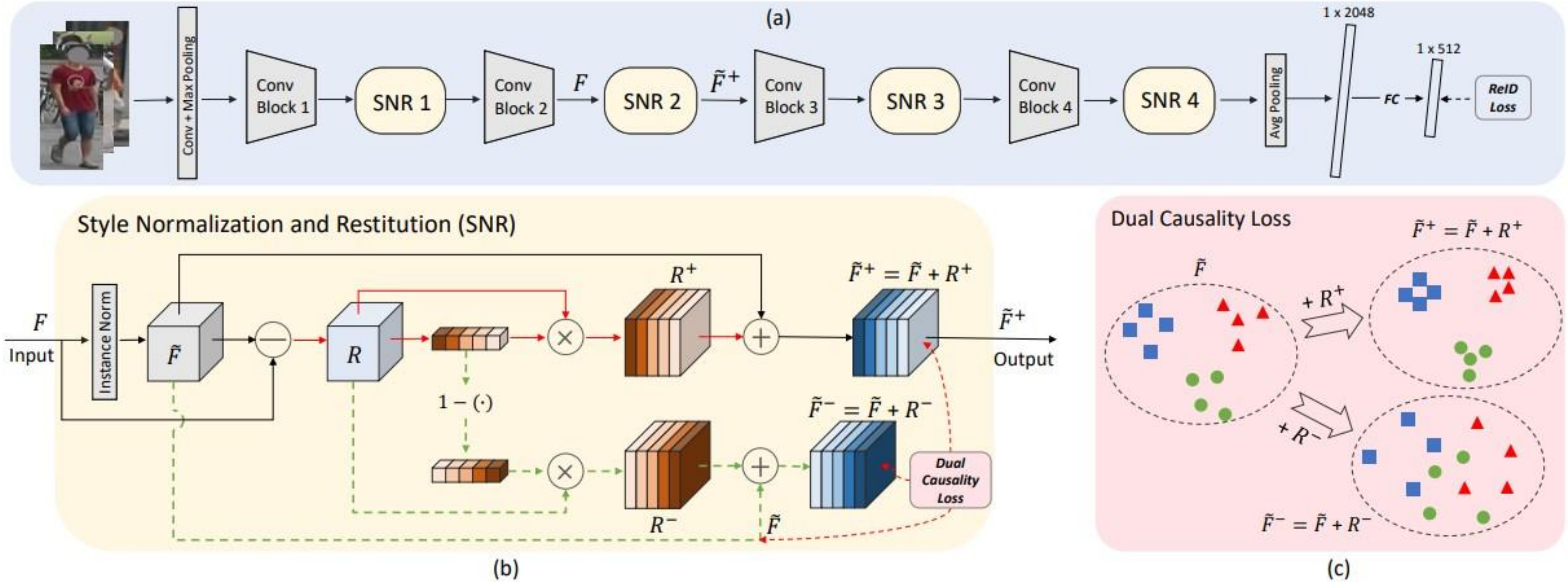
- Use **style normalization** to alleviate style variations
- This results in the **loss of some discriminative information**
- Propose to further restitute such information



Method

Style Normalization and Restitution (SNR)

- SNR can be used as a plug-and-play module and added after each convolution block
- Use ResNet-50 as backbone

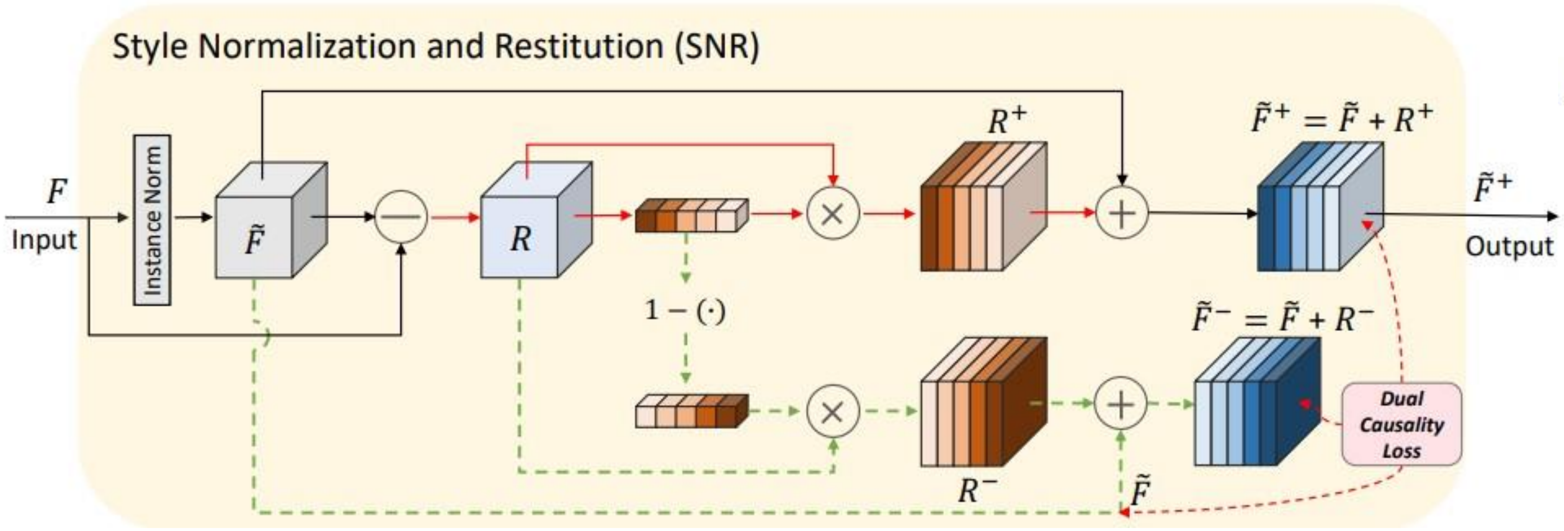


Method

Style Normalization and Restitution (SNR)

- Instance normalization performs style normalization, but inevitably removes some discriminative information

F : feature map



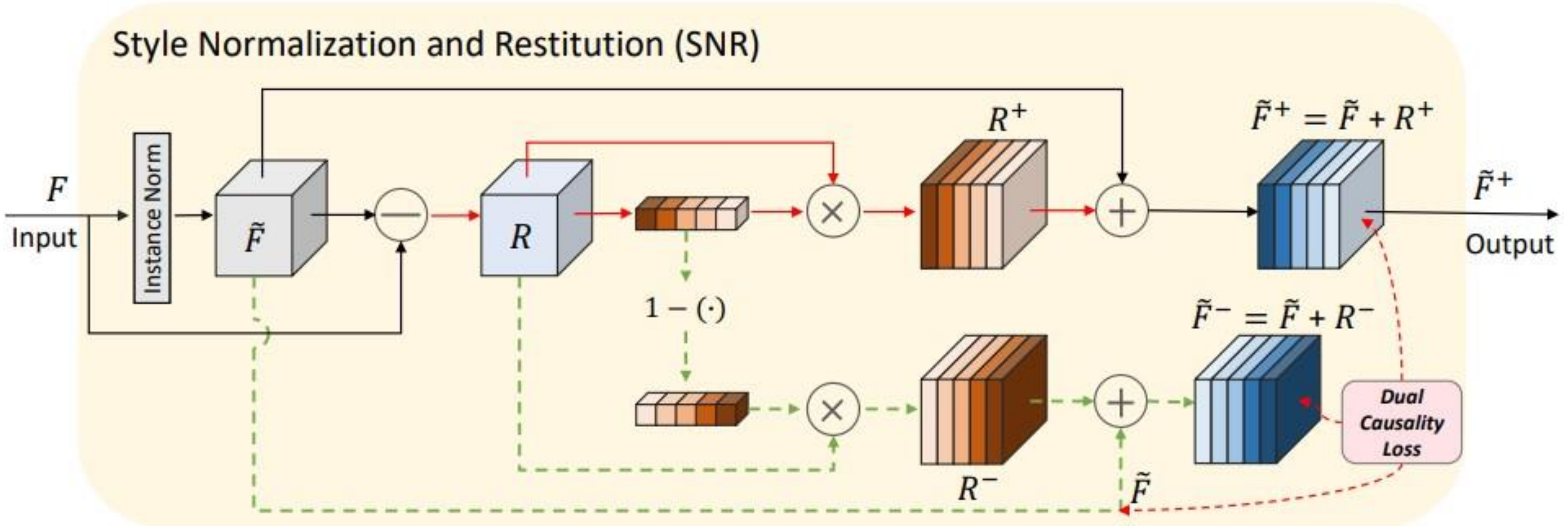
$$\widetilde{F} = \mathrm{IN}(F) = \gamma(\frac{F - \mu(F)}{\sigma(F)}) + \beta$$

Method

Style Normalization and Restitution (SNR)

- IN reduces style discrepancy and boosts generalization capability
- Define **residual feature** R as difference between the *original input feature and the style normalized feature*
- With given R, disentangle it into identity-relevant feature (+) and identity-irrelevant feature (-)

F : feature map

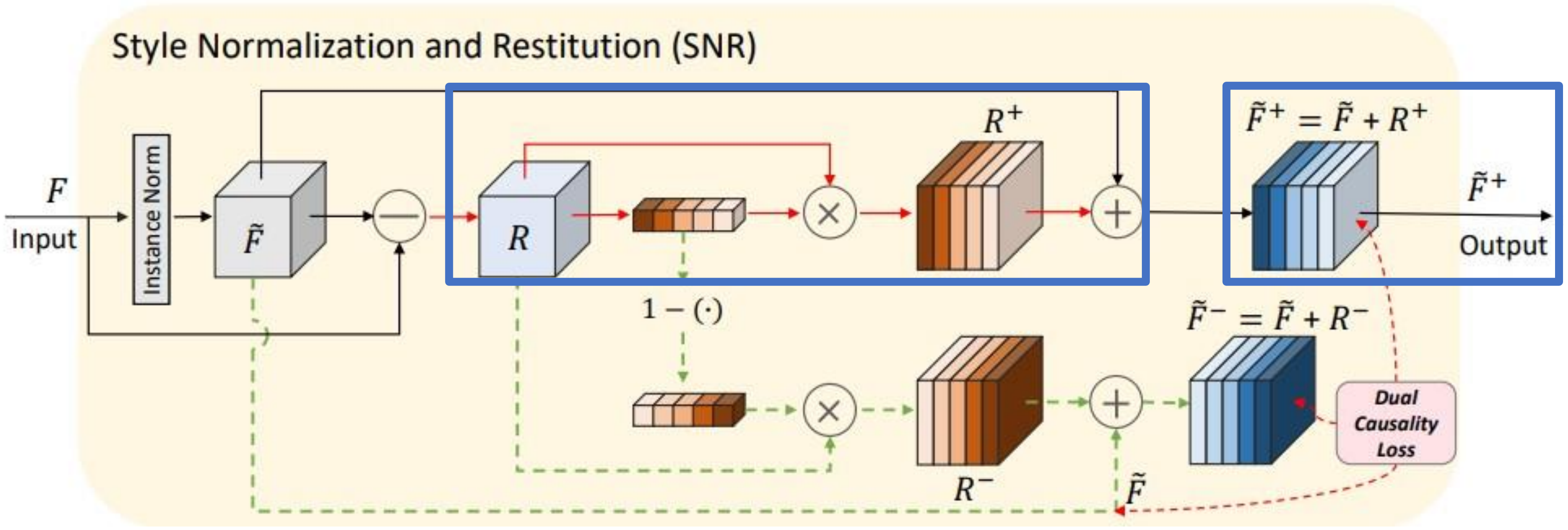


Method

Style Normalization and Restitution (SNR)

- With learned channel attention vector $\mathbf{a} = [a_1, a_2, \cdots, a_c] \in \mathbb{R}^c$
- Final output feature is obtained with adding identity-relevant feature to the style normalized feature

F : feature map

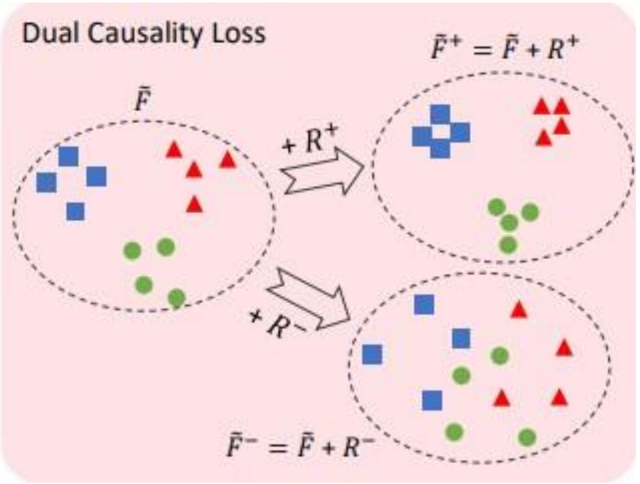


$$\mathbf{a} = g(R) = \sigma(\mathrm{W}_2\delta(\mathrm{W}_1 pool(R)))$$

Method

Style Normalization and Restitution (SNR)

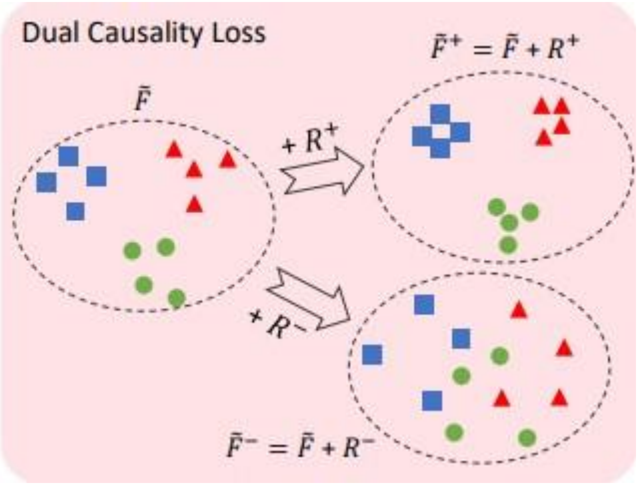
- After restituting the identity-relevant feature(+) to the normalized feature, the feature becomes more discriminative
- On the other, identity-irrelevant feature (-) should become less discriminative
- Dual causality loss helps to do it



Method

Style Normalization and Restitution (SNR)

- Consists of clarification loss(+) and destruction loss(-) $\mathcal{L}^{+}_{SNR} + \mathcal{L}^{-}_{SNR}$
- Calculate the distances between samples on a spatially average pooled feature
- **A**nchor, **P**ositive, **N**egative



$$\mathcal{L}^{+}_{SNR} = Softplus(d(\widetilde{\mathbf{f}}^{+}_{a}, \widetilde{\mathbf{f}}^{+}_{p}) - d(\widetilde{\mathbf{f}}_{a}, \widetilde{\mathbf{f}}_{p})) + Softplus(d(\widetilde{\mathbf{f}}_{a}, \widetilde{\mathbf{f}}_{n}) - d(\widetilde{\mathbf{f}}^{+}_{a}, \widetilde{\mathbf{f}}^{+}_{n}))$$

Clarification

$$Softplus(\cdot) = ln(1 + exp(\cdot))$$

$$\mathcal{L}^{-}_{SNR} = Softplus(d(\widetilde{\mathbf{f}}_{a}, \widetilde{\mathbf{f}}_{p}) - d(\widetilde{\mathbf{f}}^{-}_{a}, \widetilde{\mathbf{f}}^{-}_{p})) + Softplus(d(\widetilde{\mathbf{f}}^{-}_{a}, \widetilde{\mathbf{f}}^{-}_{n}) - d(\widetilde{\mathbf{f}}_{a}, \widetilde{\mathbf{f}}_{n})).$$

Destruction

Experiments

Source	Method	Target: Market1501 mAP	Target: Market1501 Rank-1	Target: Duke mAP	Target: Duke Rank-1	Target: PRID mAP	Target: PRID Rank-1	Target: GRID mAP	Target: GRID Rank-1	Target: VIPeR mAP	Target: VIPeR Rank-1	Target: iLIDs mAP	Target: iLIDs Rank-1
Market1501 (M)	Baseline	82.8	93.2	19.8	35.3	13.7	6.0	25.8	16.0	37.6	28.5	61.5	53.3
	Baseline-A-IN	75.3	89.8	24.1	42.7	33.9	21.0	35.6	27.2	38.1	29.1	64.2	55.0
	Baseline-IBN	81.1	92.2	21.5	39.2	19.1	12.0	27.5	19.2	32.1	23.4	58.3	48.3
	Baseline-A-SN	83.2	93.9	20.1	38.0	35.4	25.0	29.0	22.0	32.2	23.4	53.4	43.3
	Baseline-IN	79.5	90.9	25.1	44.9	35.0	25.0	35.7	27.8	35.1	27.5	64.0	54.2
	Baseline-SNR (Ours)	**84.7**	**94.4**	**33.6**	**55.1**	**42.2**	**30.0**	**36.7**	**29.0**	**42.3**	**32.3**	**65.6**	**56.7**
Duke (D)	Baseline	21.8	48.3	71.2	83.4	15.7	11.0	14.5	8.8	37.0	26.9	68.3	58.3
	Baseline-A-IN	26.5	56.0	64.5	78.9	38.6	29.0	19.6	13.6	35.1	27.2	67.4	56.7
	Baseline-IBN	24.6	52.5	69.5	81.4	27.4	19.0	19.9	12.0	32.8	23.4	63.5	61.7
	Baseline-A-SN	25.3	55.0	**73.0**	**85.9**	41.4	32.0	18.8	12.8	31.3	24.1	64.8	63.3
	Baseline-IN	27.2	58.5	68.9	80.4	40.5	27.0	20.3	13.2	34.6	26.3	70.6	65.0
	Baseline-SNR (Ours)	**33.9**	**66.7**	72.9	84.4	**45.4**	**35.0**	**35.3**	**26.0**	**41.2**	**32.6**	**79.3**	**68.7**
M + D + CUHK03 + MSMT17	Baseline	72.4	88.7	70.1	83.8	39.0	28.0	29.6	20.8	52.1	41.5	89.0	85.0
	Baseline-SNR (Ours)	**82.3**	**93.4**	**73.2**	**85.5**	**60.0**	**49.0**	**41.3**	**30.4**	**65.0**	**55.1**	**91.9**	**87.0**

Experiments

Ablation

(a) Study on the dual causality loss constraint.

Method	M⟶D mAP	M⟶D Rank-1	D⟶M mAP	D⟶M Rank-1
Baseline	19.8	35.3	21.8	48.3
SNR w/o $\mathcal{L}_{SNR}$	26.1	45.0	29.2	57.4
SNR w/o $\mathcal{L}_{SNR}^{+}$	28.8	48.9	30.2	59.8
SNR w/o $\mathcal{L}_{SNR}^{-}$	28.0	48.1	30.3	59.1
SNR	**33.6**	**55.1**	**33.9**	**66.7**

(b) Study on which stage to add SNR.

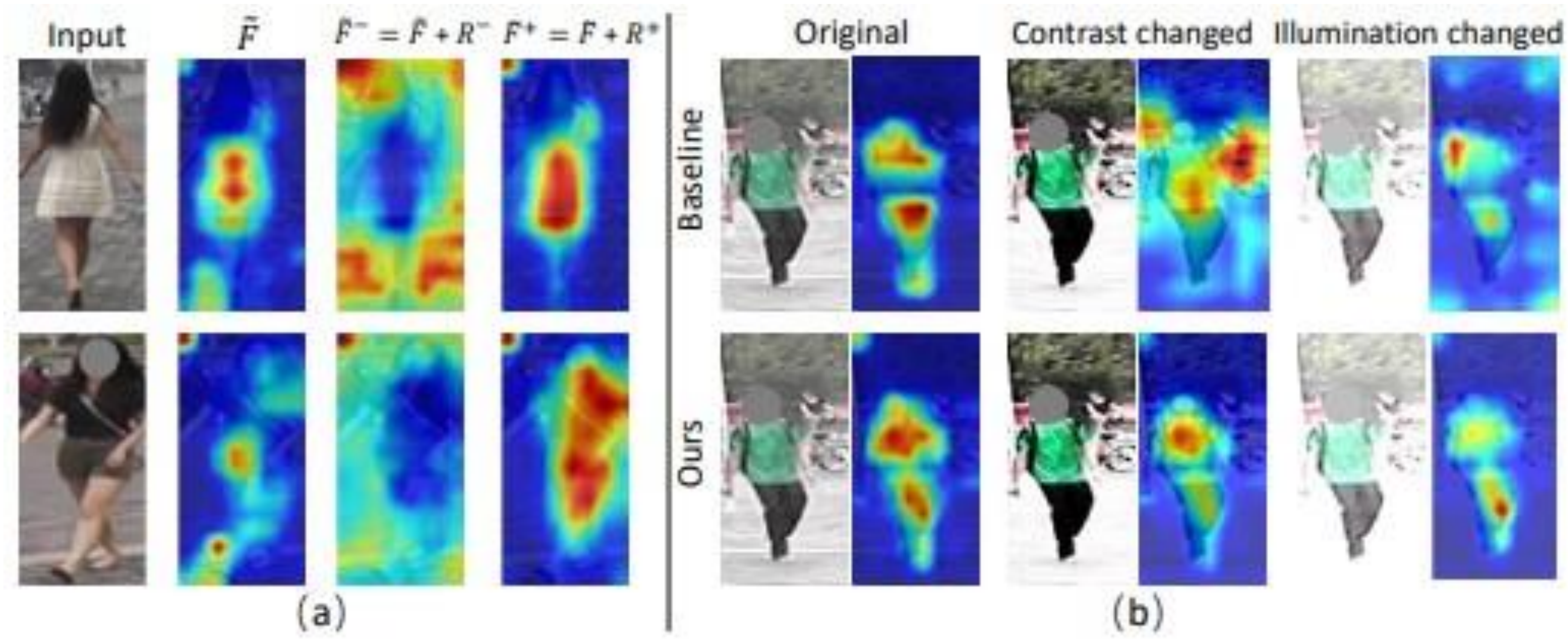
Method	M⟶D mAP	M⟶D Rank-1	D⟶M mAP	D⟶M Rank-1
Baseline	19.8	35.3	21.8	48.3
stage-1	23.7	42.8	27.6	57.7
stage-2	24.0	44.4	28.6	58.8
stage-3	26.4	46.3	29.5	60.7
stage-4	26.2	45.8	29.4	59.7
stages-all	**33.6**	**55.1**	**33.9**	**66.7**

(c) Disentanglement designs in SNR.

Method	M⟶D mAP	M⟶D Rank-1	D⟶M mAP	D⟶M Rank-1
Baseline	19.8	35.3	21.8	48.3
SNR_{conv}	29.7	51.1	29.4	61.7
$\mathrm{SNR}_{g(\cdot)^2}$	31.2	52.9	31.0	63.8
SNR	**33.6**	**55.1**	**33.9**	**66.7**

Experiments

Visualizations



Conclusion

Conclusion

- Propose a generalizable person ReID framework to enable

- SNR module exploit the Instance Normalization that filters out the interference from style variations, and restitute the identity relevant features that are discarded by IN

- To disentangle the identity-relevant and –irrelevant features, they further design a dual causality loss constraint

Thank you !