RLE: A Unified Perspective of Data Augmentation for Cross-Spectral Re-identification

Lei Tan[1], Yukang Zhang[1], Keke Han[1], Pingyang Dai[1(✉)],

Yongjian Wu[2], Rongrong Ji

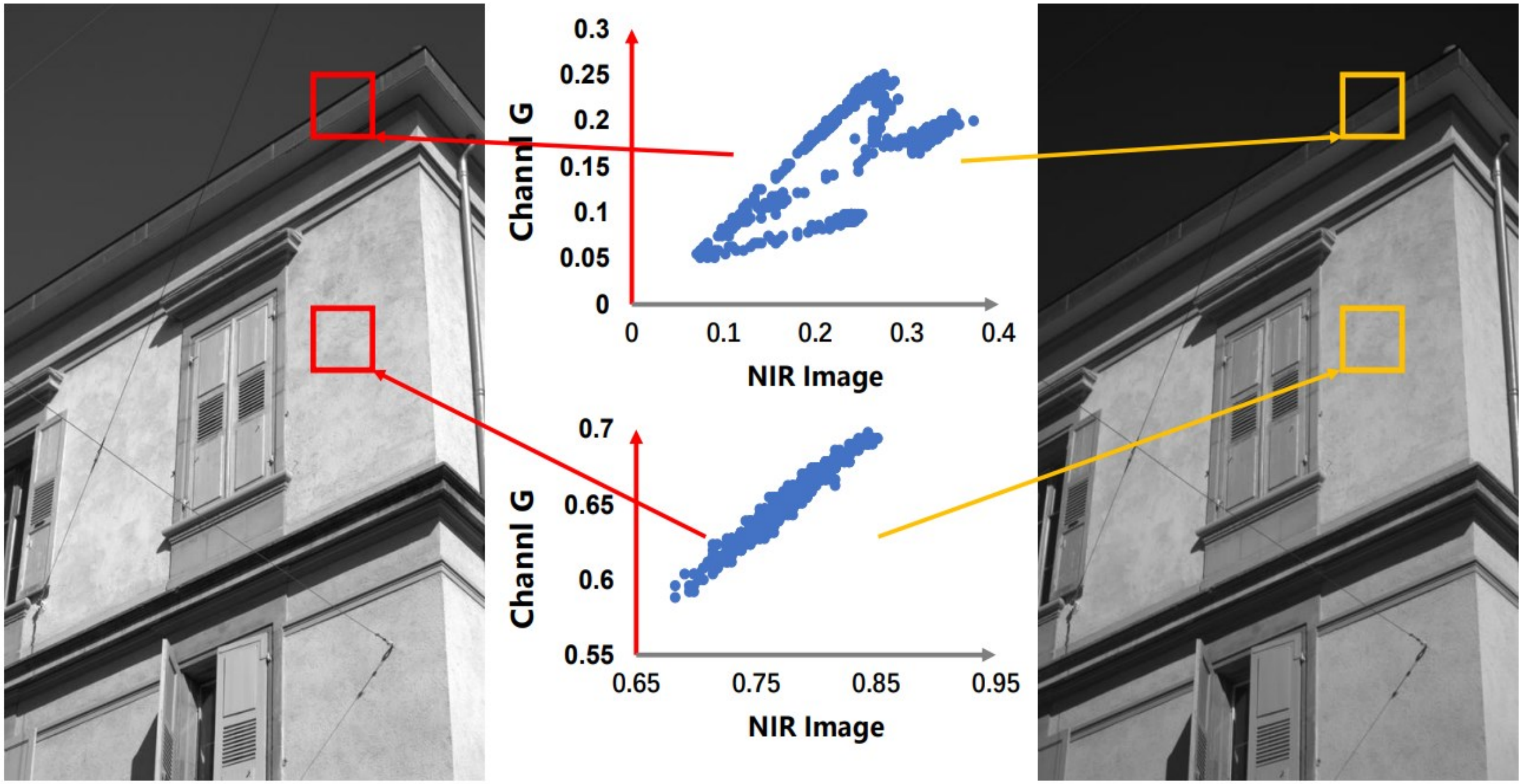
[1]Key Laboratory of Multimedia Trusted Perception and Efficient Computing,

Ministry of Education of China, Xiamen University, China.

[2]Tencent Youtu Lab, China.

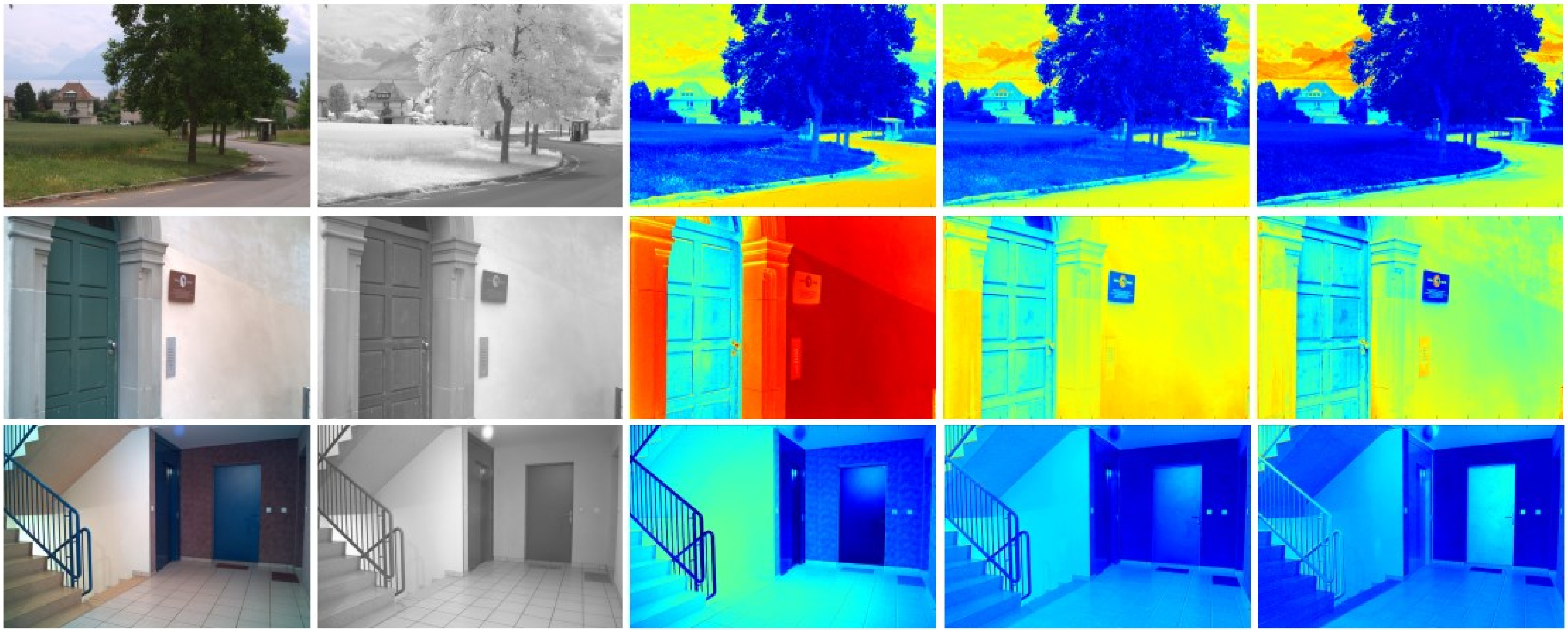
Code: https://github.com/stone96123/RLE

Observation



- Under the same illumination, the cross-spectral transformation could be described as a linear transformation in a material similar surface. Still, in the whole image level, the transformation is nonlinear due to the diversity of materials.

Observation



VIS NIR R-NIR G-NIR B-NIR

With Lambertain model:

$$\frac{\rho_N(x,y)}{\rho_G(x,y)} = \frac{\omega_N \beta_N(x,y) \int_{\lambda_N} F_N(\lambda) S(\lambda,x,y) Q_N(\lambda) d\lambda}{\underbrace{\omega_G \beta_G(x,y)}_{\text{Constant}} \int_{\lambda_G} \underbrace{F_G(\lambda)}_{\text{Light source}} \underbrace{S(\lambda,x,y)}_{\text{Material}} \underbrace{Q_G(\lambda)}_{\text{Camera}} d\lambda}.$$

The linear factor is only influence by the **Material**

- Data augmentation for cross-spectral re-identification is formed to achieve non-linear transformations with different distinct local linear factors, thus encouraging the network to be robust to such a transformation.

Contributions

- As an effort to model the transformation behind the modality discrepancy in the cross-spectral Re-ID task, we discover that the cross-spectral **modality discrepancy mainly comes from different local linear transformations caused by the diversity of materials**. Based on this observation, we further categorize the cross-spectral data augmentation strategies into moderate and radical transformations under a unified perspective.

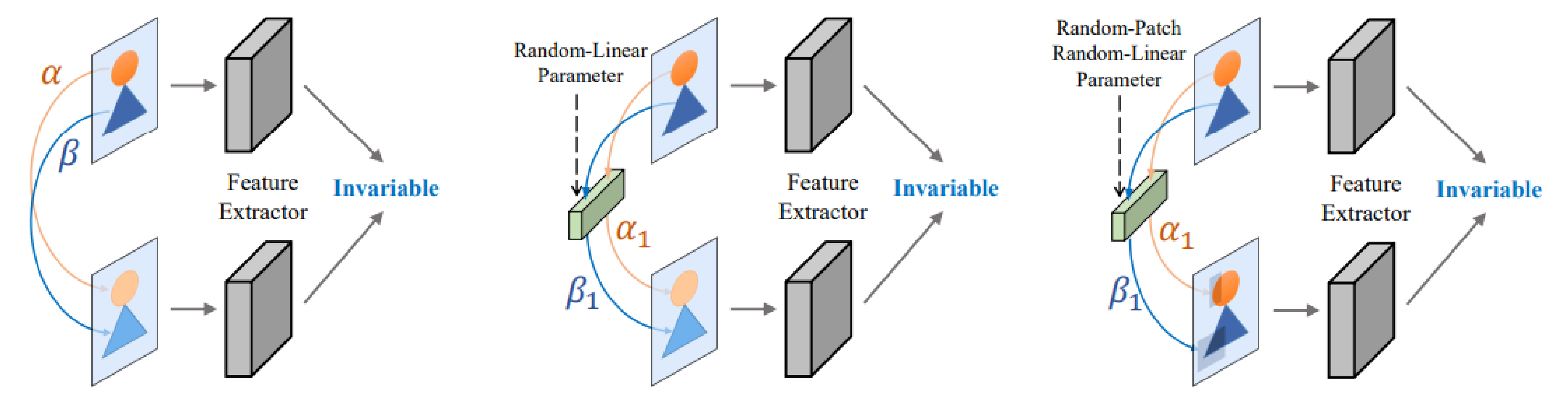
- By extending the observation, we propose a **Random Linear Enhancement** (RLE) strategy, which includes **Moderate Random Linear Enhancement** (MRLE) and **Radical Random Linear Enhancement** (RRLE). The RLE effectively takes advantage of the aforementioned unified perspective and embeds it in a controllable linear transformation.

- Extensive experiments on cross-spectral re-identification datasets demonstrate the effectiveness and superior ability of the proposed RLE, which can **boost performance under various scenarios**.

Method

Random Linear Enhancement

- **Overview**



(a) Definite linear transformation on Definite image patch. (b) Random linear transformation on Definite image patch. (c) Random linear transformation on Random image patch.

- **Moderate RLE**
  - **Moderate Random Linear Enhancement is designed to provide diverse image transformations that satisfy the original linear correlations under constrained conditions**

Moderate Transformation:

$$I_{mt} = \lambda_r I_r + \lambda_g I_g + \lambda_b I_b,$$
$$s.t. \quad \lambda_r + \lambda_g + \lambda_b = 1,$$

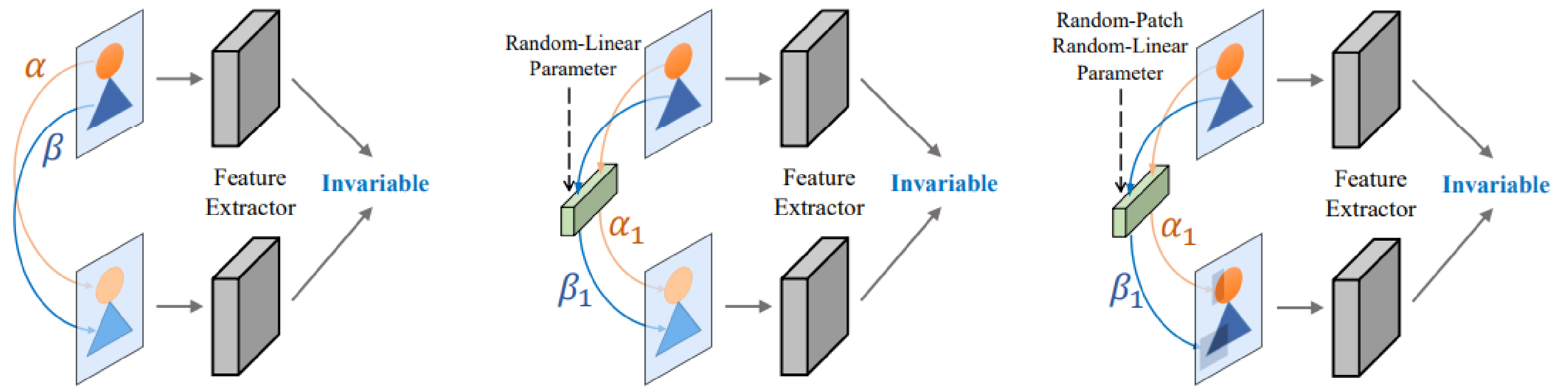
Moderate RLE:

$$I_{mt} = \lambda_r I_r + \lambda_g I_g + \lambda_b I_b,$$
$$with \quad \lambda_r, \lambda_g, \lambda_b \sim Beta(\beta_m, \beta_m),$$
$$s.t. \quad \lambda_r + \lambda_g + \lambda_b = 1.$$

Method

Random Linear Enhancement

- **Overview**



(a) Definite linear transformation on Definite image patch. (b) Random linear transformation on Definite image patch. (c) Random linear transformation on Random image patch.

- **Radical RLE**
  - **Radical Random Linear Enhancement seeks to generate local linear transformations directly without relying on external information**

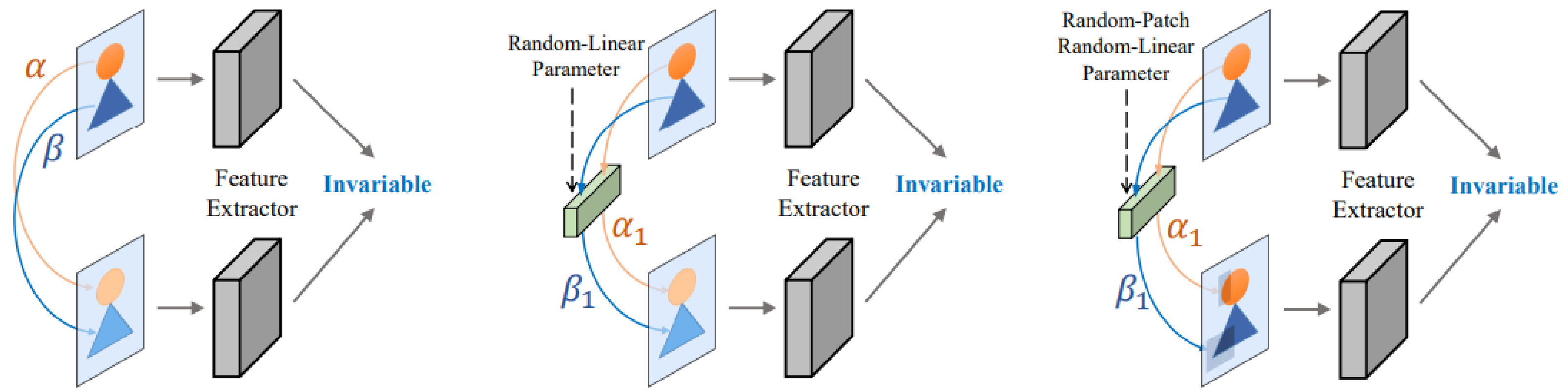
Radical RLE:

$$
\begin{aligned}
&I_{mt} = \lambda_r I_r + \lambda_g I_g + \lambda_b I_b, \\
&with \quad \lambda_r, \lambda_g, \lambda_b \sim Beta(\beta_m, \beta_m), \\
&s.t. \quad \lambda_r + \lambda_g + \lambda_b = 1.
\end{aligned}
$$

Method

Random Linear Enhancement

- **Overview**



(a) Definite linear transformation on Definite image patch. (b) Random linear transformation on Definite image patch. (c) Random linear transformation on Random image patch.

- **Radical RLE**
  - **Radical Random Linear Enhancement seeks to generate local linear transformations directly without relying on external information**

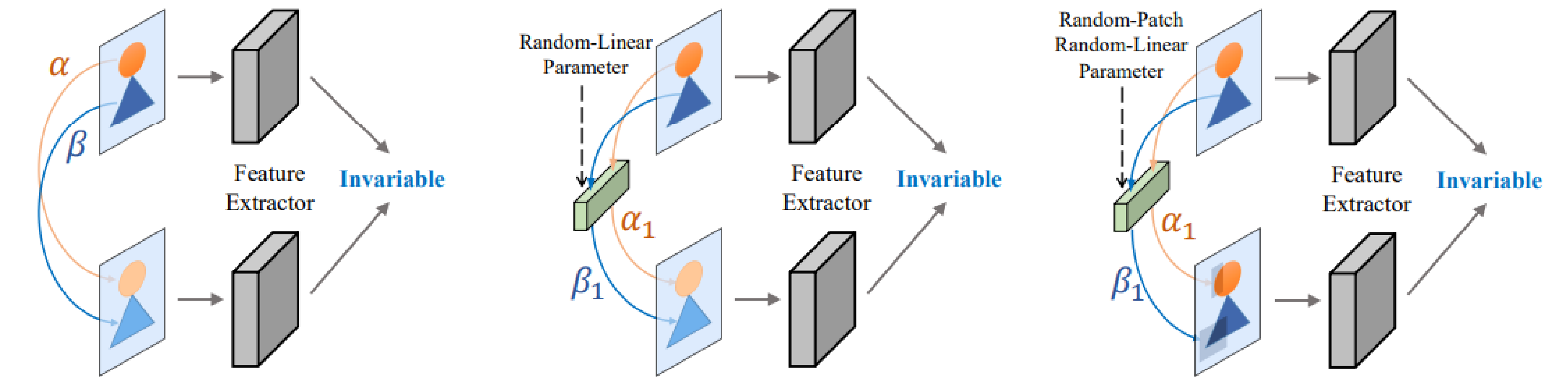
Radical RLE:

$$
\begin{aligned}
&I_{mt} = \lambda_r I_r + \lambda_g I_g + \lambda_b I_b, \\
&with \quad \lambda_r, \lambda_g, \lambda_b \sim Beta(\beta_m, \beta_m), \\
&s.t. \quad \lambda_r + \lambda_g + \lambda_b = 1.
\end{aligned}
$$

Method

Random Linear Enhancement

- **Overview**



(a) Definite linear transformation on Definite image patch. (b) Random linear transformation on Definite image patch. (c) Random linear transformation on Random image patch.

- **Radical RLE**
  - **Radical Random Linear Enhancement seeks to generate local linear transformations directly without relying on external information**

Radical RLE:

$$
\begin{aligned}
&I_{mt} = \lambda_r I_r + \lambda_g I_g + \lambda_b I_b,\\
&with \quad \lambda_r, \lambda_g, \lambda_b \sim Beta(\beta_m, \beta_m),\\
&s.t. \quad \lambda_r + \lambda_g + \lambda_b = 1.
\end{aligned}
$$

Experiments

Table 1: Ablation study of different data augmentation strategies on the cross-spectral re-identification task. 'Gray' denotes the grayscale transformation, 'RC' refers to the random channel selection. 'MRLE' indicates the moderate random linear enhancement. 'RE' refers to the random erasing, and 'RRLE' means the radical random linear enhancement.

Setting	*All Search*						*Indoor Search*					
	R-1	R-5	R-10	R-20	*m*AP	*m*INP	R-1	R-5	R-10	R-20	*m*AP	*m*INP
Moderate Transformation												
Baseline	64.5	88.1	94.2	98.1	62.9	50.4	70.0	91.7	96.6	99.1	75.1	71.1
Baseline+Gray	66.7	89.4	95.2	98.7	64.2	50.8	72.0	93.9	97.8	99.6	77.1	73.0
Baseline+RC	68.3	90.6	95.6	98.6	65.3	51.9	72.7	93.9	97.6	99.7	77.7	73.7
Baseline+RC+Gray	68.6	90.7	96.0	98.8	64.9	52.3	74.3	94.1	98.1	99.6	78.8	75.1
Baseline+MRLE	**70.2**	**91.6**	**96.5**	**99.0**	**67.0**	**53.5**	**75.5**	**95.2**	**98.2**	**99.7**	**79.7**	**75.9**
Radical Transformation												
Baseline+RE	71.0	91.4	96.3	99.1	69.5	57.4	78.5	95.8	98.7	99.8	82.1	78.4
Baseline+RRLE	72.0	92.4	97.2	99.4	69.1	56.3	77.0	96.4	**99.1**	**99.9**	81.4	77.7
Baseline+RE+RRLE	**74.2**	**93.0**	**97.4**	**99.5**	**71.8**	**60.4**	**81.7**	**96.7**	**99.1**	**99.9**	**84.5**	**81.2**
Mixed Transformation												
Baseline+CAJ [11]	73.5	92.9	97.4	99.4	69.4	55.4	80.7	96.1	98.6	99.8	83.5	79.8
Baseline+RLE+RE	**75.4**	**93.5**	**97.7**	**99.6**	**72.4**	**60.9**	**84.7**	**97.9**	**99.3**	**99.9**	**87.0**	**83.7**

(a) **Performance under different** β_m. Compared to the uniform distribution, a U-shaped beta distribution works better in MRLE.

β_m	R-1	*m*AP	*m*INP
1.0	67.9	65.3	52.2
0.5	67.8	65.1	51.8
0.4	68.3	65.6	52.3
0.3	**70.2**	**67.0**	**53.5**
0.2	67.9	66.0	52.6

(b) **Performance under different** β_r. Compared to the uniform distribution, a U-shaped beta distribution works better in MRLE.

β_r	R-1	*m*AP	*m*INP
0.5	72.5	70.7	59.3
0.4	**74.2**	**71.8**	**60.4**
0.3	73.2	71.3	60.2
0.2	73.7	71.7	60.2

(c) **Performance under different** t_{min}. Using too small t_{min} will introduce excessive noise and lead to performance degradation.

t_{min}	R-1	*m*AP	*m*INP
0.3	72.9	70.9	58.2
0.2	73.8	71.3	59.0
0.1	**74.2**	**71.8**	**60.4**
0.01	73.8	71.7	59.8
0.001	73.6	71.3	59.5

Table 4: Applicability of our opposed RLE to other methods on the SYSU-MM01 dataset.

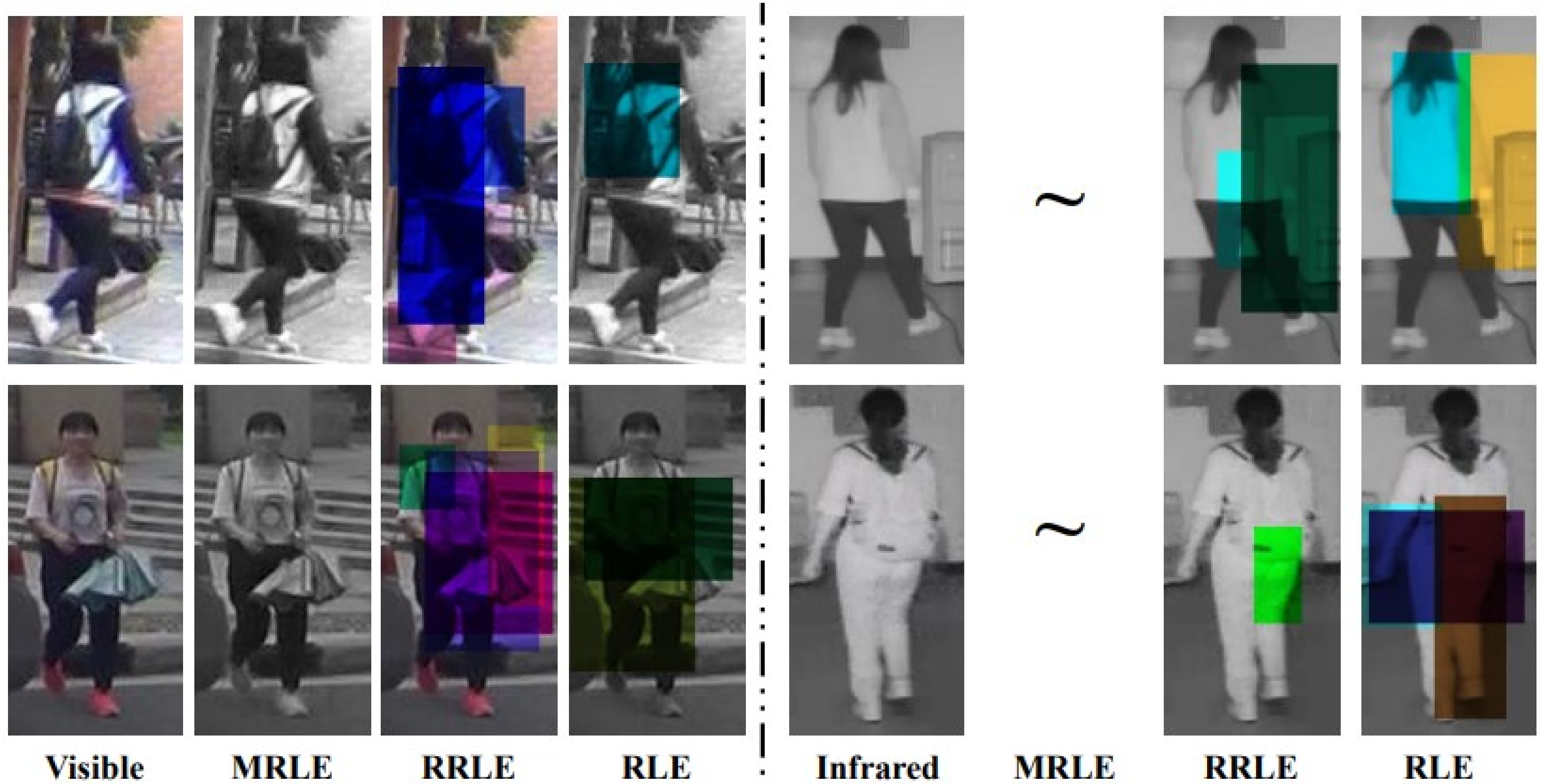
Setting	*All Search*		*Indoor Search*	
	R-1	*m*AP	R-1	*m*AP
DEEN [39]	74.7	71.8	80.3	83.3
+Ours	**76.2** (+1.5)	**73.0** (+1.2)	**83.2** (+2.9)	**85.3** (+2.0)
ViT-B [49]	66.0	63.1	69.9	75.1
+Ours	**70.2**(+4.2)	**66.7**(+3.6)	**71.9**(+2.0)	**76.4**(+1.3)

Experiments

Table 2: Comparisons between the proposed method and some state-of-the-art methods on the SYSU-MM01 and RegDB datasets.

Methods	SYSU-MM01						RegDB					
	All Search			*Indoor Search*			*VIS to IR*			*IR to VIS*		
	R-1	R-10	mAP	R-1	R-10	mAP	R-1	R-10	mAP	R-1	R-10	mAP
BDTR[40]	17.0	55.4	19.7	-	-	-	33.6	58.6	32.8	32.9	58.5	32.0
D^{2}RL[5]	28.9	70.6	29.2	-	-	-	43.4	66.1	44.1	-	-	-
Hi-CMD[41]	34.9	77.6	35.9	-	-	-	70.9	86.4	66.0	-	-	-
AlignGAN[10]	42.4	85.0	40.7	45.9	87.6	54.3	57.9	-	53.6	56.3	-	53.4
DDAG[36]	54.8	90.4	53.0	61.0	94.1	68.0	69.3	86.2	63.5	68.1	85.2	61.8
LbA[42]	55.4	-	54.1	58.5	-	66.3	74.2	-	67.6	67.5	-	72.4
NFS[43]	56.9	91.3	55.5	62.8	96.5	69.8	80.5	91.6	72.1	78.0	90.5	69.8
CM-NAS[44]	60.8	92.1	58.9	68.0	94.8	52.4	82.8	95.1	79.3	81.7	94.1	77.6
MCLNet[37]	65.4	93.3	62.0	72.6	97.0	76.6	80.3	92.7	73.1	75.9	90.9	69.5
FMCNet[45]	66.3	-	62.5	68.2	-	74.1	89.1	-	84.4	88.4	-	83.9
SMCL[46]	67.4	92.9	61.8	68.8	96.6	75.6	83.9	-	79.8	83.1	-	78.6
DART[20]	68.7	96.4	66.3	72.5	97.8	78.2	83.6	-	75.7	82.0	-	73.8
CAJ[11]	69.9	95.7	66.9	76.3	97.9	80.4	85.0	95.5	79.1	84.8	95.3	77.8
MPANet[47]	70.6	96.2	68.2	76.7	98.2	81.0	82.8	-	80.7	83.7	-	80.9
MMN [25]	70.6	96.2	66.9	76.2	97.2	79.6	91.6	97.7	84.1	87.5	96.0	80.5
MAUM [48]	71.7	-	68.8	77.0	-	81.9	87.9	-	-	87.0	-	84.3
CAJ+ [12]	71.5	96.2	68.2	78.4	98.4	82.0	85.7	95.5	79.7	84.9	95.9	78.6
DEEN [39]	74.7	97.6	71.8	80.3	99.0	83.3	91.1	97.8	85.1	89.5	96.8	83.4
Ours	**75.4**	**97.7**	**72.4**	**84.7**	**99.3**	**87.0**	**92.8**	**97.9**	**88.6**	**91.0**	**97.5**	**86.6**

Experiments



- MRLE provides an efficient way to provide **diverse transformations from multi-spectral images to single-spectral images**, while the RRLE gets rid of the dependence on the multiple spectral images and **makes such a linear transformation directly on the local part**.

Thanks for your attention!