DONGYUXU

Dec 1 · 3 min read · Listen

Reading Notes: Inter-Task Association for Cross-Resolution Person Re-Identification

> *This paper proposes a training regularisation called Inter-Task Association Critic (INTACT) to accelerate the integrated model (Supre-Resolution and Reid) training.*

MOTIVATION: The multi-task joint learning framework cascades SR and Reid to address the cross-resolution tasks is dramatically more difficult gradients backpropagation.

Below shows the plain cascade model's objective function:

$$\mathcal{L}_{\text{sr}} = \mathcal{L}_{\text{MSE}} + \lambda_g \mathcal{L}_{\text{gan}} + \lambda_c \mathcal{L}_{\text{id}}$$

λg and λc are weight parameters

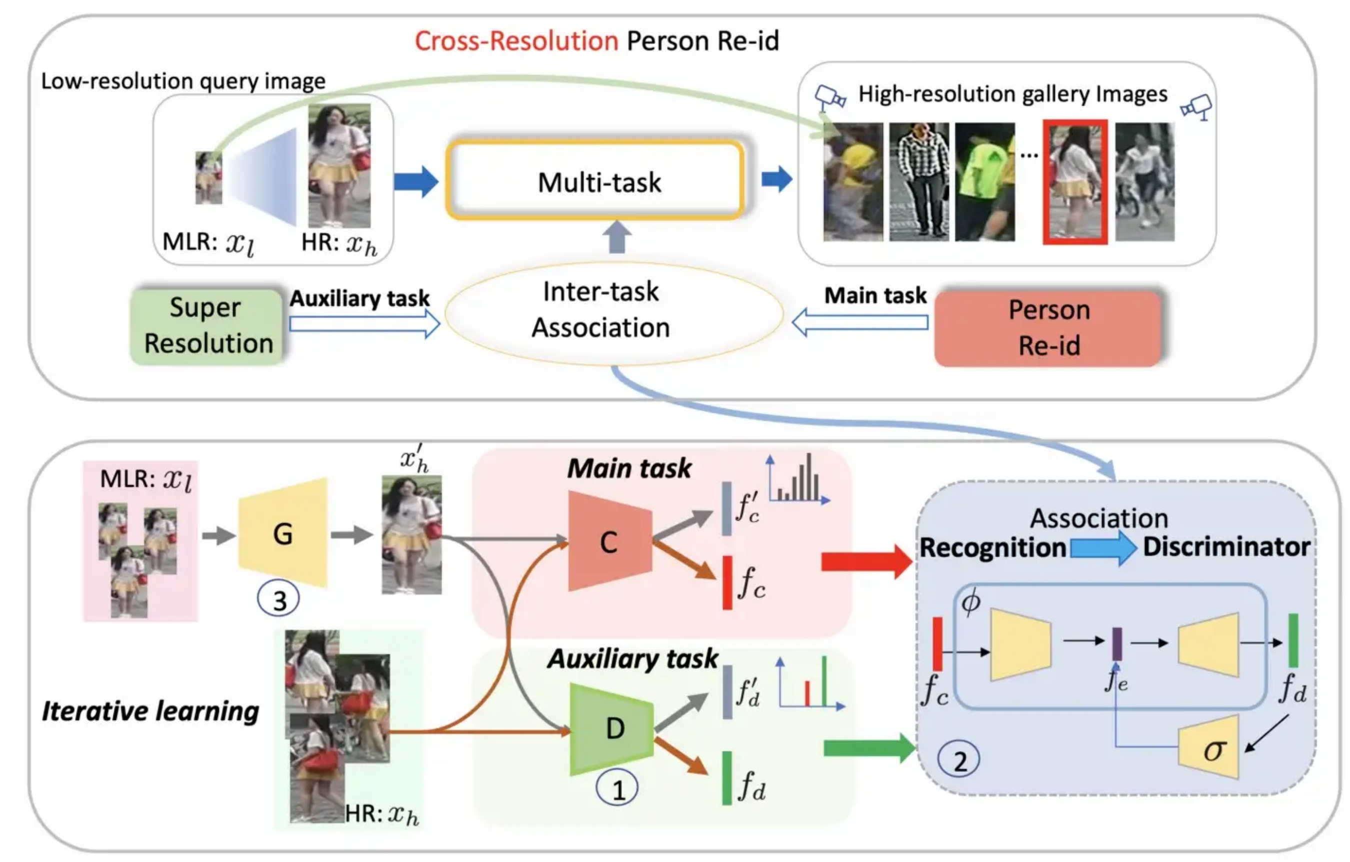
where

$$\mathcal{L}_{\text{gan}} = \mathbb{E}_{x_h}[\log D(x_h)] + \mathbb{E}_{x_l}[\log (1 - D(G(x_l))]$$

x_h and x_l stand for high-resolution and low-resolution images

$$\mathcal{L}_{\text{MSE}} = \|x_h - G(x_l)\|_2^2.$$

METHODOLOGY: Use a dedicated network \phi to represent the intrinsic association between the SR and Reid.



- Part I: Association Learning

$$\mathcal{L}_{\text{intact}} = ||\phi(\boldsymbol{f}_{\text{c}}) - \boldsymbol{f}_{\text{d}}||_2^2$$

f_c is the identity classification feature extracted from the Reid and f_d is the discriminator feature

To optimize the parameters of /epi

$$\mathcal{L}_{\mathrm{e}} = ||\sigma(\boldsymbol{f}_{\mathrm{d}}) - \boldsymbol{f}_{\mathrm{e}}||_2^2$$

σ is a transform of the target fd

Add an additional bridging constraint to manipulate the optimizing direction

$$\mathcal{L}_{\text{intact-e}} = \mathcal{L}_{\text{intact}} + \mathcal{L}_{\mathrm{e}}$$

The association model /epi and the bridging model σ are jointly learned

- Part II: Association Regularisation

$$\mathcal{L}_{\mathrm{dis}} = ||\phi(\boldsymbol{f}_{\mathrm{c}}') - \boldsymbol{f}_{\mathrm{d}}'||_2^2$$

f_c^’ and f_d^’ are the identity and discriminator of the SR model

The association network /epi is fixed to serve as an external critic

Brief summary: Use a model to learn the association between the real identity classification feature and the discriminator feature. After that, fixed the model parameters as a constraint to train the SR model(Generator), with the hope that the identity representation from SR could be more like Reid.

TRAINING

Algorithm 1 INTACT model training

Input: Training data $\mathcal{D} = \{x_l, x_h\}$ with identity labels Y.
Output: A person image super-resolution (SR) model.
Initialisation: Training a standard person re-id model with HR images and the identity labels.
Alternating training (frozen one, and update the others):
for $i = 1$ **to** *iter* **do**
(1) Update the discriminator with the GAN loss (Eq. (2));
(2) Update the association network ϕ (Eq. (7));
(3) Update the generator (SR model) with the SR objective loss (Eq. (4)) and distillation loss (Eq. (8)).
end for

$$G^* = \arg\min_G \max_D \mathcal{L}_{\text{gan}}. \tag{2}$$

$$\mathcal{L}_{\text{intact-e}} = \mathcal{L}_{\text{intact}} + \mathcal{L}_{\text{e}} \tag{7}$$

$$\mathcal{L}_{\text{sr}} = \mathcal{L}_{\text{MSE}} + \lambda_g \mathcal{L}_{\text{gan}} + \lambda_c \mathcal{L}_{\text{id}} \tag{4}$$

$$\mathcal{L}_{\text{dis}} = ||\phi(\boldsymbol{f}'_{\text{c}}) - \boldsymbol{f}'_{\text{d}}||_2^2 \tag{8}$$

EXPERIMENTS

Table 1. Cross-resolution person re-id performance (%). Bold and underlined numbers indicate top two results, respectively.

Model	MLR-Market-1501			MLR-CUHK03			MLR-VIPeR			MLR-DukeMTMC-reID			CAVIAR		
	Rank1	Rank5	Rank10	Rank1	Rank5	Rank10	Rank1	Rank5	Rank10	Rank1	Rank5	Rank10	Rank1	Rank5	Rank10
CamStyle [51]	74.5	88.6	93.0	69.1	89.6	93.9	34.4	56.8	66.6	64.0	78.1	84.4	32.1	72.3	85.9
FD-GAN [12]	79.6	91.6	93.5	73.4	93.8	97.9	39.1	62.1	72.5	67.5	82.0	85.3	33.5	71.4	86.5
SLD^{2}L [17]	-	-	-	-	-	-	20.3	44.0	62.0	-	-	-	18.4	44.8	61.2
SING [16]	74.4	87.8	91.6	67.7	90.7	94.7	33.5	57.0	66.5	65.2	80.1	84.8	33.5	72.7	89.0
CSR-GAN [40]	76.4	88.5	91.9	71.3	92.1	97.4	37.2	62.3	71.6	67.6	81.4	85.1	34.7	72.5	87.4
JUDEA [25]	-	-	-	26.2	58.0	73.4	26.0	55.1	69.2	-	-	-	22.0	60.1	80.8
SDF [39]	-	-	-	22.2	48.0	64.0	9.3	38.1	52.4	-	-	-	14.3	37.5	62.5
RAIN [7]	-	-	-	78.9	97.3	98.7	42.5	68.3	79.6	-	-	-	42.0	77.3	89.6
CAD [26]	83.7	92.7	95.8	82.1	97.4	**98.8**	43.1	68.2	77.5	75.6	86.7	89.6	42.8	76.2	91.5
INTACT (Ours)	**88.1**	**95.0**	**96.9**	**86.4**	**97.4**	98.5	**46.2**	**73.1**	**81.6**	**81.2**	**90.1**	**92.8**	**44.0**	**81.8**	**93.9**

Table 3. Evaluating INTACT's loss components on MLR-Market-1501. MSE: pixel-wise content loss, ID: identity classification loss (Eq. (3)), Association: our association loss (Eq. (7) & (8)).

Supervision	Rank1	Rank5	Rank10
MSE+ID	83.7	93.0	95.6
MSE+ID+GAN	84.7	93.9	96.1
MSE+ID+GAN+Association	**88.1**	**95.0**	**96.9**

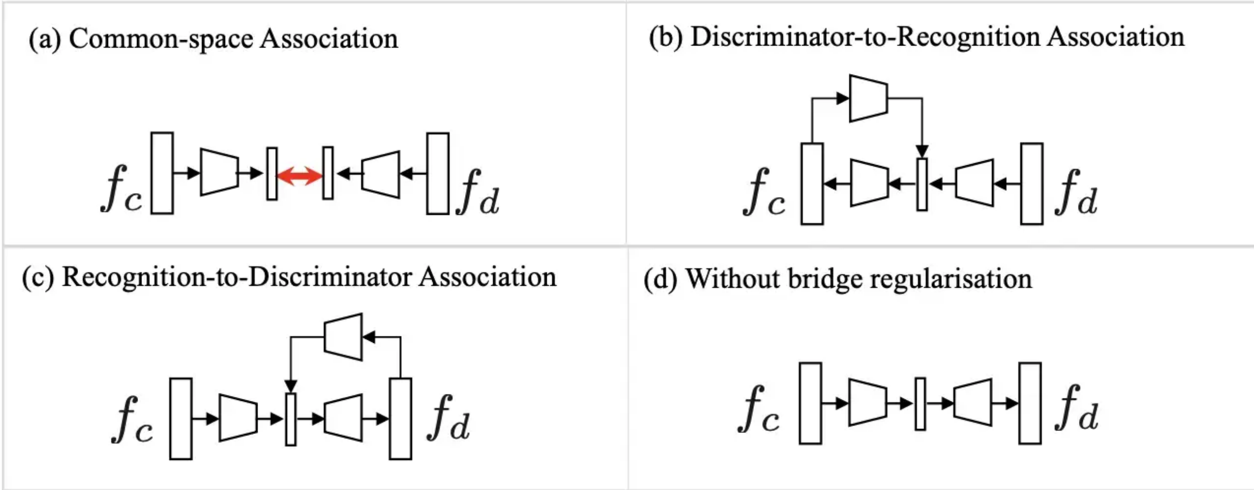


Figure 5. Schematics of different association designs.

Association Space	Rank1	Rank5	Rank10
Common Space (a)	84.3	94.0	95.3
D-to-R (b)	83.4	93.5	95.0
R-to-D (c, ours)	**88.1**	**95.0**	**96.9**

Table 5. Effect of the bridge constraint (Eq. (6)).

Bridge constraint	Rank1	Rank5	Rank10
W/O (Fig. 5 (d))	84.3	93.5	95.8
W (Fig. 5 (c))	**88.1**	**95.0**	**96.9**