

Mitigate Catastrophic Remembering via Continual Knowledge Purification for Noisy Lifelong Person Re-Identification

Anonymous Authors

In our supplementary materials, we provide more theoretical analysis and implementation details of our proposed approach. Besides, we also provide more quantitative and qualitative results of our methods in comparison with existing methods.

In summary, the supplementary materials primarily contain:

- Theoretical analysis of our Cluster-based Label Scoring (CLS) strategy.
- More model implementation details about our anti-forgetting loss functions \mathcal{L}_{wKD} .
- Relevant ReID datasets and arrangement in our proposed NLRReID benchmark.
- Quantitative comparison with more state-of-the-art LReID and LNL methods under different label noise ratios including the 0% noise data.
- Qualitative analysis of the person retrieval capacity of our model in comparison with the state-of-the-art competitor.

1 THEORETICAL ANALYSIS OF CLS

In our Cluster-based Label Scoring (CLS) strategy, given N_t image-label pairs $\{(x_i, y_i)\}_{i=1}^{N_t}$, the features are extracted by the model and the clustering strategy is adopted to obtain clusters (the outliers are also gathered as an extra cluster). Then, according to the clustering results, each image x_i is assigned a cluster-aware label $\tilde{y}_i \in \{1, 2, \dots, N_c\}$ where N_c is the cluster number.

Subsequently, we generate one-hot embedding $\mathbf{l}_i \in \mathbb{R}^{N_c}$ for each x_i according to \tilde{y}_i . Therefore, \mathbf{l}_i satisfies:

$$1 = \sum_{j=1}^{N_c} \mathbf{l}_i^{(j)}, \quad (1)$$

where $\mathbf{l}_i^{(j)}$ is the j -th element of \mathbf{l}_i , and for arbitrary $j \in \{1, 2, \dots, N_c\}$, $0 \leq \mathbf{l}_i^{(j)} \leq 1$.

Besides, according to Equation (1) in the main paper, we obtain the annotation-aware average cluster label

$$\bar{\mathbf{l}}_i = \frac{1}{n_t^i} \sum_{j=1}^{N_t} \delta(y_i, y_j) \mathbf{l}_j, \quad (2)$$

where $0 \leq \bar{\mathbf{l}}_i^{(k)}$ ($k \in \{1, 2, \dots, N_c\}$) and $\delta(y_i, y_j)$ is a sign function that outputs 1 and 0 when $y_i = y_j$ and $y_i \neq y_j$ respectively. n_t^i indicates the number of images that are annotated with the same label with x_i . Specifically, n_t^i is obtained by

$$n_t^i = \sum_{j=1}^{N_t} \delta(y_i, y_j), \quad (3)$$

where $n_t^i \geq 1$ since $\delta(y_i, y_j) \geq 0$ and $\delta(y_i, y_j) = 1$ when $i = j$.

Then, the label distance d_i calculation process in Equation (2) of the main paper could be derived as follows:

$$\begin{aligned} d_i &= \|\mathbf{l}_i - \bar{\mathbf{l}}_i\|_2^2 \\ &= \|\mathbf{l}_i - \frac{1}{n_t^i} \sum_{j=1}^{N_t} \delta(y_i, y_j) \mathbf{l}_j\|_2^2 \\ &= \|\frac{n_t^i - 1}{n_t^i} \mathbf{l}_i - \frac{1}{n_t^i} \sum_{j \neq i}^{N_t} \delta(y_i, y_j) \mathbf{l}_j\|_2^2 \\ &= \sum_{k=1}^{N_c} (\frac{n_t^i - 1}{n_t^i} \mathbf{l}_i^{(k)} - \frac{1}{n_t^i} \sum_{j \neq i}^{N_t} \delta(y_i, y_j) \mathbf{l}_j^{(k)})^2 \\ &\leq \sum_{k=1}^{N_c} (\frac{n_t^i - 1}{n_t^i} \mathbf{l}_i^{(k)})^2 + \sum_{k=1}^{N_c} (\frac{1}{n_t^i} \sum_{j \neq i}^{N_t} \delta(y_i, y_j) \mathbf{l}_j^{(k)})^2 \\ &= (\frac{n_t^i - 1}{n_t^i})^2 \sum_{k=1}^{N_c} (\mathbf{l}_i^{(k)})^2 + (\frac{1}{n_t^i})^2 \sum_{k=1}^{N_c} (\sum_{j \neq i}^{N_t} \delta(y_i, y_j) \mathbf{l}_j^{(k)})^2 \quad (4) \\ &\leq (\frac{n_t^i - 1}{n_t^i})^2 (\sum_{k=1}^{N_c} \mathbf{l}_i^{(k)})^2 + (\frac{1}{n_t^i})^2 (\sum_{j \neq i}^{N_t} \sum_{k=1}^{N_c} \delta(y_i, y_j) \mathbf{l}_j^{(k)})^2 \\ &= (\frac{n_t^i - 1}{n_t^i})^2 + (\frac{1}{n_t^i})^2 (\sum_{j \neq i}^{N_t} \delta(y_i, y_j))^2 \\ &= (\frac{n_t^i - 1}{n_t^i})^2 + (\frac{1}{n_t^i})^2 (n_t^i - 1)^2 \\ &= 2(\frac{n_t^i - 1}{n_t^i})^2 \\ &< 2 \end{aligned}$$

Therefore, $d_i \in [0, 2)$ and simply $s_i = (2 - d_i)/2$ could convert the label distance d_i into label confidence $s_i \in [0, 1]$ (the d_i of the outliers during clustering are set to 2, thereby allowing the possibility of $s_i = 0$).

2 MODEL IMPLEMENTATION DETAILS

We have shown a complete data process pipeline of our Continual Knowledge Purification (CKP) in the main paper, here we primarily show the implementation details of the modified knowledge distillation loss \mathcal{L}_{wKD} in our work built upon the CORE and LSTKC baseline (CORE[†]).

Specifically, given a batch of B images, we utilize the EKF algorithm in our main paper to obtain w_t^o for each image and generate the subset \mathcal{B}_o with B_o images by removing the images with $w_t^o < T_o$. Thus the features extracted by M_t can be represented as $\mathbf{F}_t^o \in \mathbb{R}^{B_o \times d}$. The pair-wise similarity matrix $\mathbf{S}_t^o \in \mathbb{R}^{B_o \times B_o}$ is calculated by

$$\mathbf{S}_t^o = \mathbf{F}_t^o \cdot (\mathbf{F}_t^o)^\top. \quad (5)$$

Table 1: The statistics of datasets used in the NLReID benchmark. ‘-’ indicates the datasets only serve as test domains.

Type	Datasets Name	Original Identities			NLReID Identities		
		Train	Query	Gallery	Train	Query	Gallery
Seen	CUHK03 [7]	767	700	700	500	700	700
	Market-1501 [24]	751	750	751	500	751	751
	DukeMTMC-ReID [13]	702	702	1110	500	702	1110
	CUHK-SYSU [16]	942	2900	2900	500	2900	2900
	MSMT17-V2 [15]	1041	3060	3060	500	3060	3060
Unseen	i-LIDS [1]	243	60	60	-	60	60
	VIPeR [3]	316	316	316	-	316	316
	GRID [10]	125	125	126	-	125	126
	PRID [4]	100	100	649	-	100	649
	CUHK01 [6]	485	486	486	-	486	486
	CUHK02 [5]	1677	239	239	-	239	239
	SenseReID [23]	1718	521	1718	-	521	1718

Table 2: Results under the Random Noise. [†] indicates the state-of-the-art LNL method is combined with the latest anti-forgetting strategy of LSTKC.

Metric	Method	Market-1501			CUHK-SYSU			DukeMTMC			MSMT17			CUHK03			Seen-Avg			Unseen-Avg		
		10%	20%	30%	10%	20%	30%	10%	20%	30%	10%	20%	30%	10%	20%	30%	10%	20%	30%	10%	20%	30%
mAP	PTKP [2]	38.4	25.2	18.1	74.3	67.4	65.0	29.2	20.0	14.2	9.1	5.8	3.8	23.4	11.9	6.4	34.9	26.1	21.5	44.4	35.2	30.3
	SELF [19]	32.3	24.0	15.6	71.5	69.1	62.8	34.1	21.2	11.8	10.6	7.2	3.8	44.4	30.0	16.2	38.6	30.3	22.1	49.8	41.2	31.8
	PNET [21]	23.2	19.0	12.7	69.2	65.5	58.5	21.0	14.6	8.6	6.8	4.7	2.5	38.2	29.4	19.8	31.7	26.6	20.4	43.0	36.4	30.0
	PNET [†] [21]	46.8	30.7	23.5	78.6	69.7	66.2	18.3	6.9	5.6	5.4	2.1	1.6	12.4	6.4	5.0	32.3	23.2	20.4	40.3	30.8	26.3
	SELF [†] [19]	46.7	41.7	33.3	79.8	77.7	71.3	44.4	29.9	12.8	17.7	8.2	3.6	34.5	20.2	10.3	44.7	35.5	26.2	52.9	43.5	34.4
	CKP(Ours)	48.7	44.5	42.2	80.8	80.3	78.6	47.3	44.4	42.1	18.1	16.4	14.6	42.0	36.3	33.5	47.4	44.4	42.2	56.0	51.4	50.4
R@1	PTKP [2]	65.5	51.6	41.9	77.7	71.9	69.6	48.8	36.1	27.9	26.7	20.0	14.5	24.2	12.1	5.5	48.6	38.3	31.9	38.0	30.2	25.0
	SELF [19]	58.7	49.6	36.8	74.7	73.2	66.5	53.3	37.0	25.1	29.5	21.7	13.5	47.9	30.1	15.9	52.8	42.3	31.6	43.7	35.8	26.5
	PNET [21]	47.2	41.3	30.1	72.9	69.1	62.3	36.5	27.3	18.8	19.9	14.6	8.8	40.9	30.8	19.4	43.5	36.6	27.9	37.1	30.7	24.7
	PNET [†] [21]	70.4	55.8	46.5	81.4	73.0	69.6	31.7	16.2	13.6	15.3	6.9	5.4	12.3	5.4	4.4	42.2	31.5	27.9	33.9	25.1	20.5
	SELF [†] [19]	71.7	66.6	59.0	82.4	80.4	74.9	63.5	48.3	26.3	40.6	22.6	12.1	35.6	19.7	9.0	58.8	47.5	36.2	46.0	37.4	29.1
	CKP(Ours)	71.8	68.1	66.9	83.2	83.0	81.0	64.7	62.1	58.9	40.1	37.5	34.7	42.4	37.1	34.1	60.4	57.6	55.1	49.5	44.5	43.4

Then, the pair-wise relation matrix $R_t^o \in \mathbb{R}^{B_o \times B_o}$ is calculated by

$$R_t^o[i, j] = \frac{\exp(S_t^o[i, j]/\tau_o)}{\sum_{k=1}^{B_o} \exp(S_t^o[i, k]/\tau_o)}, \quad (6)$$

where τ_o is a temperature hyper-parameter to scale the values of R_t^o . τ_o is set to 0.1 by default following [17]. Therefore the weighted knowledge distillation loss \mathcal{L}_{wKD} is calculated by

$$\mathcal{L}_{wKD} = \frac{1}{B_o} \sum_{i=1}^{B_o} \mathcal{L}_{KL}(R_{t-1}^o[i, :] || R_t^o[i, :]), \quad (7)$$

where R_{t-1}^o is obtained by utilizing \mathcal{B}_o and M_{t-1} following the above process. \mathcal{L}_{KL} is the Kullback Leibler (KL) divergence function.

3 DATASETS DETAILS

We build the Noisy LReID benchmark (NLReID) based on LReID [11]. A total of 12 existing ReID datasets are adopted (Market1501 [24], DukeMTMC-reID [13], CUHK-SYSU [16], MSMT17-V2 [15], CUHK03

[7], CUHK01 [6], CUHK02 [5], VIPeR [3], PRID [4], i-LIDS [1], GRID [10], and SenseReID [23]). The detailed statistics of these datasets are shown in Table 1. Among these datasets, CUHK-SYSU is special because it was initially developed for the person search task. To transform it into a ReID dataset, the ground-truth person bounding box annotation is used to crop individual-level images, and a subset in which each identity contains at least 4 bounding boxes is selected. In addition, to mitigate the data imbalance between datasets [11, 12], 500 identities of each dataset are selected to form the NLReID benchmark. Specifically, in Table 1, the ‘Original Identities’ indicates the identity numbers in the original datasets, and ‘NLReID Identities’ indicates the selected identity numbers in the NLReID benchmark.

4 QUANTITATIVE RESULTS

In this section, we implement more LReID and LNL methods [2, 20, 21] under different levels of random and patterned label noise in Table 2 and Table 3 respectively. We also report the results of different methods under 0% noise data (*i.e.*, clean data) in Table 4.

Table 3: Results under the Patterned Noise. [†] indicates the state-of-the-art LNL method is combined with the latest anti-forgetting strategy of LSTKC.

Metric	Method	Market-1501			CUHK-SYSU			DukeMTMC			MSMT17			CUHK03			Seen-Avg			Unseen-Avg		
		10%	20%	30%	10%	20%	30%	10%	20%	30%	10%	20%	30%	10%	20%	30%	10%	20%	30%	10%	20%	30%
mAP	PTKP [2]	39.1	29.3	24.7	75.0	69.6	66.7	30.7	22.5	18.1	9.2	6.6	5.5	24.1	16.1	10.2	35.6	28.8	25.0	45.3	36.6	34.4
	SELF [19]	33.5	29.6	23.6	71.3	71.2	67.8	33.2	28.5	18.4	11.4	9.3	6.3	44.0	35.4	24.7	38.7	34.8	28.2	48.7	44.0	38.3
	PNET [21]	23.9	19.2	18.3	69.4	66.6	63.3	19.2	16.1	12.4	6.5	5.0	3.6	37.8	31.4	24.6	31.4	27.7	24.4	43.1	38.7	34.7
	PNET [†] [21]	49.3	35.8	28.5	80.0	73.2	68.2	24.6	9.8	6.6	6.6	2.5	1.7	12.3	7.6	8.5	34.6	25.8	22.7	43.4	32.5	30.8
	SELF [†] [19]	46.0	41.4	42.0	79.1	78.3	76.6	45.0	40.1	33.9	18.0	13.8	8.4	38.7	27.0	17.5	45.4	40.1	35.7	52.8	47.1	44.4
	CKP(Ours)	50.1	46.9	44.1	81.0	79.9	78.9	47.2	45.2	43.1	18.3	17.1	15.7	41.9	39.9	36.6	47.7	45.8	43.7	57.3	54.4	51.1
R@1	PTKP [2]	64.6	55.2	51.2	78.4	73.3	71.0	49.2	40.8	34.8	26.5	21.4	18.3	24.2	15.6	9.1	48.6	41.3	36.9	40.0	31.0	29.0
	SELF [19]	61.0	56.0	48.1	74.3	74.3	71.9	53.9	47.1	33.7	30.6	27.5	19.5	44.6	37.0	25.7	52.9	48.4	39.8	41.6	38.3	32.9
	PNET [21]	48.7	41.1	40.2	72.9	70.3	67.2	34.9	29.6	24.0	19.7	15.5	12.1	38.9	31.5	25.4	43.0	37.6	33.8	37.8	33.1	29.1
	PNET [†] [21]	72.2	59.8	51.4	82.6	76.5	72.0	40.6	19.8	15.1	18.4	8.3	6.0	11.1	7.2	7.3	45.0	34.3	30.4	36.8	27.1	25.0
	SELF [†] [19]	71.0	66.1	67.0	81.3	81.1	79.9	64.9	59.3	53.5	41.5	34.7	23.9	40.9	26.9	16.9	59.9	53.6	48.2	45.9	40.8	38.5
	CKP(Ours)	73.2	70.8	67.5	82.7	82.3	81.4	65.8	62.5	61.2	41.0	38.9	36.7	42.7	40.8	37.0	61.1	59.1	56.8	50.1	47.5	44.6

Table 4: Results under the No noise. [†] indicates the LNL method is combined with the anti-forgetting strategy of the LReID method LSTKC.

Noise	Type	Method	Reference	Market1501 mAP R@1	CUHK-SYSU mAP R@1	DukeMTMC mAP R@1	MSMT17 mAP R@1	CUHK03 mAP R@1	Seen-Avg mAP R@1	Unseen-Avg mAP R@1
0%	-	Joint-Train	-	75.3 90.1	84.5 86.0	66.9 81.6	31.6 57.1	58.5 61.4	63.4 75.2	55.2 48.2
0%	LReID	LwF [9]	T-PAMI'17	56.3 77.1	72.9 75.1	29.6 46.5	6.0 16.6	36.1 37.5	40.2 50.6	47.2 42.6
		PatchKD [14]	ACM MM'22	68.5 85.7	75.6 78.6	33.8 50.4	6.5 17.0	34.1 36.8	43.7 53.7	49.1 45.4
		KRKC [22]	AAAI'23	39.8 65.3	75.9 78.3	35.1 53.2	7.8 21.0	57.9 59.7	43.3 55.5	53.6 48.2
		PTKP [2]	AAAI'22	54.8 75.7	81.5 83.9	44.7 62.6	15.9 36.3	47.1 48.4	48.8 61.4	56.7 50.1
		LSTKC [17]	AAAI'24	54.7 76.0	81.1 83.4	49.4 66.2	20.0 43.2	44.7 46.5	50.0 63.1	57.0 49.9
	LNL	PNET [21]	TIFS'20	27.3 51.2	59.6 62.9	27.5 46.1	7.7 20.9	49.2 50.6	34.3 46.3	42.7 34.6
		SELF [19]	TIP'21	33.5 57.9	66.3 68.8	34.4 53.7	10.7 27.3	50.5 52.1	39.1 52.0	46.4 39.0
		CORE [19]	TIP'21	37.8 64.3	76.9 79.5	38.1 57.8	14.9 36.2	46.5 48.9	42.8 57.3	53.6 46.6
		DICS [8]	CVPR'23	23.2 44.6	47.1 48.8	21.8 39.1	5.2 14.5	37.5 37.9	26.9 37.0	38.6 31.7
	Noisy LReID	PNET [†] [21]	TIFS'20	56.5 77.5	81.0 83.0	50.4 67.5	17.8 39.8	41.0 41.5	49.3 61.9	56.7 49.3
		SELF [†] [19]	TIP'21	52.4 74.3	80.1 82.4	48.7 66.2	19.3 42.3	45.7 47.7	49.2 62.6	56.7 49.8
		CORE [†] [19]	TIP'21	52.4 75.0	81.1 82.9	48.1 65.6	19.7 42.7	45.3 46.8	49.3 62.6	55.8 48.0
		DICS [†] [8]	CVPR'23	48.2 72.4	81.5 83.9	42.8 61.9	18.1 42.4	28.8 29.7	43.9 58.1	53.1 46.9
		LCNL [†] [18]	IJCV'24	53.5 75.4	80.2 82.0	49.9 68.1	18.7 41.2	43.9 45.4	49.2 62.4	56.0 48.7
		CKP(Ours)	-	53.8 76.0	81.2 83.0	49.7 67.0	18.4 40.8	44.1 45.8	49.4 62.5	58.0 51.0

Comparison with More Existing Methods under Noisy Data As shown in Table 2 and Table 3, we additionally compare with LReID method PTKP[2] and LNL method PNET [21], SELF [21]. The results show that our model achieves superior performance on each individual dataset and outperforms the existing methods significantly across different noise ratios under both random and patterned noise. In particular, compared to the best competitor SELF[†], our CKP achieves **16.0%/18.9%**, **16.0%/14.3%** average mAP/R@1 improvement in seen and unseen domains respectively under 30% random noise in Table 2. Besides, **8.0%/8.6%**, **6.7%/6.1%** average mAP/R@1 improvement is achieved in seen and unseen domains

respectively under 30% patterned noise in Table 3, compared to the SELF[†]. Therefore, our model is more robust to noise data compared to existing methods. This is because our CKP continually rectifies the wrong labels and generates high-quality training data for new knowledge learning, and our knowledge selection algorithm mitigates the erroneous old knowledge accumulation so that some erroneous knowledge can be rectified by the new model.

Comparison with Existing Methods under Clean Data As shown in Table 4, our CKP achieves comparable performance with state-of-the-art method LSTKC [17] on seen domains, with only a

marginal decline of **0.6%/0.6%** average mAP/R@1. The slight inferiority in the seen domain results from the inherent trade-off in our data purification and erroneous knowledge filtering designs, which unavoidably lead to a slight loss of new knowledge. However, our method exhibits state-of-the-art **generalization** performance on unseen domains, surpassing LSTKC by **1.0%** in average mAP and **1.1%** in average R@1. This is because our data purification and erroneous knowledge filtering designs effectively filter out challenging knowledge that is difficult for the model to learn. Consequently, this mitigates overfitting to the seen domains and enhances the generalization capacity of the model, resulting in improved performance on unseen domains.

5 VISUALIZATION OF RETRIEVAL RESULTS

In this section, we visualize the person retrieval results of our model and the state-of-the-art method CORE[†] on different seen and unseen datasets. Both models are trained with the NLReID benchmark under 30% random noise. The results of seen datasets Market-1501 (Figure 1) CUHK-SYSU (Figure 2), MSMT17-V2 (Figure 3), CUHK03 (Figure 4) are visualized to show the acquisition and anti-forgetting capability of the models. Besides, the results on unseen datasets CUHK01 (Figure 5), CUHK02 (Figure 6), SenseReID (Figure 7), and VIPR (Figure 8) are visualized to show the generalization of the models. The images in black boxes represent the input query images, and the images in the green and red boxes represent the correct and false retrievals respectively. The retrieval results are arranged in descending order, from left to right, according to the matching scores. The results show that in scenarios involving significant variations in human pose, viewpoint, and scenery, our model could achieve more accurate retrieving, demonstrating the correct knowledge acquisition and accumulation capacity of our proposed method.

REFERENCES

- [1] Home Office Scientific Development Branch. 2006. Imagery library for intelligent detection systems (i-lids). In *2006 IET Conference on Crime and Security*. IET, 445–448.
- [2] Wenhao Ge, Junlong Du, Ancong Wu, Yuqiao Xian, Ke Yan, Feiyue Huang, and Wei-Shi Zheng. 2022. Lifelong Person Re-identification by Pseudo Task Knowledge Preservation. In *AAAI*, Vol. 36. 688–696.
- [3] Douglas Gray and Hai Tao. 2008. Viewpoint invariant pedestrian recognition with an ensemble of localized features. In *ECCV*. Springer, 262–275.
- [4] Martin Hirzer, Csaba Belezai, Peter M Roth, and Horst Bischof. 2011. Person re-identification by descriptive and discriminative classification. In *Image Analysis*. Springer, 91–102.
- [5] Wei Li and Xiaogang Wang. 2013. Locally aligned feature transforms across views. In *CVPR*. IEEE, 3594–3601.
- [6] Wei Li, Rui Zhao, and Xiaogang Wang. 2012. Human Reidentification with Transferred Metric Learning. In *ACCV*. Springer, 31–44.
- [7] Wei Li, Rui Zhao, Tong Xiao, and Xiaogang Wang. 2014. DeepReID: Deep Filter Pairing Neural Network for Person Re-identification. In *CVPR*. IEEE, 152–159.
- [8] Yifan Li, Hu Han, Shiguang Shan, and Xilin Chen. 2023. Disc: Learning from noisy labels via dynamic instance-specific selection and correction. In *CVPR*. 24070–24079.
- [9] Zhizhong Li and Derek Hoiem. 2017. Learning without Forgetting. *PAMI* 40, 12 (2017), 2935–2947.
- [10] Chen Change Loy, Tao Xiang, and Shaogang Gong. 2010. Time-delayed Correlation Analysis for Multi-camera Activity Understanding. *IJCV* 90, 1 (2010), 106–129.
- [11] Nan Pu, Wei Chen, Yu Liu, Erwin M Bakker, and Michael S Lew. 2021. Lifelong Person Re-Identification via Adaptive Knowledge Accumulation. In *CVPR*. IEEE, 7897–7906.
- [12] Nan Pu, Zhun Zhong, Nicu Sebe, and Michael S Lew. 2023. A Memorizing and Generalizing Framework for Lifelong Person Re-Identification. *PAMI* (2023).
- [13] Ergys Ristani, Francesco Solera, Roger Zou, Rita Cucchiara, and Carlo Tomasi. 2016. Performance measures and a data set for multi-target, multi-camera tracking. In *ECCV*. Springer, 17–35.
- [14] Zhicheng Sun and Yadong Mu. 2022. Patch-based Knowledge Distillation for Lifelong Person Re-Identification. In *ACMMM*. 696–707.
- [15] Longhui Wei, Shiliang Zhang, Wen Gao, and Qi Tian. 2018. Person Transfer GAN to Bridge Domain Gap for Person Re-identification. In *CVPR*. IEEE, 79–88.
- [16] Tong Xiao, Shuang Li, Bochao Wang, Liang Lin, and Xiaogang Wang. 2016. End-to-end deep learning for person search. *arXiv:1604.01850* 2, 2 (2016), 4.
- [17] Kunlun Xu, Xu Zou, and Jiahuan Zhou. 2024. LSTKC: Long Short-Term Knowledge Consolidation for Lifelong Person Re-identification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 38. 16202–16210.
- [18] Mouxing Yang, Zhenyu Huang, and Xi Peng. 2024. Robust object re-identification with coupled noisy labels. *IJCV* (2024), 1–19.
- [19] Mang Ye, He Li, Bo Du, Jianbing Shen, Ling Shao, and Steven CH Hoi. 2021. Collaborative refining for person re-identification with label noise. *IEEE Transactions on Image Processing* 31 (2021), 379–391.
- [20] Mang Ye, Jianbing Shen, Gaojie Lin, Tao Xiang, Ling Shao, and Steven CH Hoi. 2021. Deep Learning for Person Re-identification: A Survey and Outlook. *PAMI* 44, 6 (2021), 2872–2893.
- [21] Mang Ye and Pong C Yuen. 2020. PurifyNet: A robust person re-identification model with noisy labels. *TIFS* 15 (2020), 2655–2666.
- [22] Chunlin Yu, Ye Shi, Zimo Liu, Shenghua Gao, and Jingya Wang. 2023. Lifelong Person Re-Identification via Knowledge Refreshing and Consolidation. In *AAAI*, Vol. 37. 3295–3303.
- [23] Haiyu Zhao, Maoqing Tian, Shuyang Sun, Jing Shao, Junjie Yan, Shuai Yi, Xiaogang Wang, and Xiaoou Tang. 2017. Spindle Net: Person Re-identification with Human Body Region Guided Feature Decomposition and Fusion. In *CVPR*. IEEE, 907–915.
- [24] Liang Zheng, Liyue Shen, Lu Tian, Shengjin Wang, Jingdong Wang, and Qi Tian. 2015. Scalable person re-identification: A benchmark. In *ICCV*. IEEE, 1116–1124.

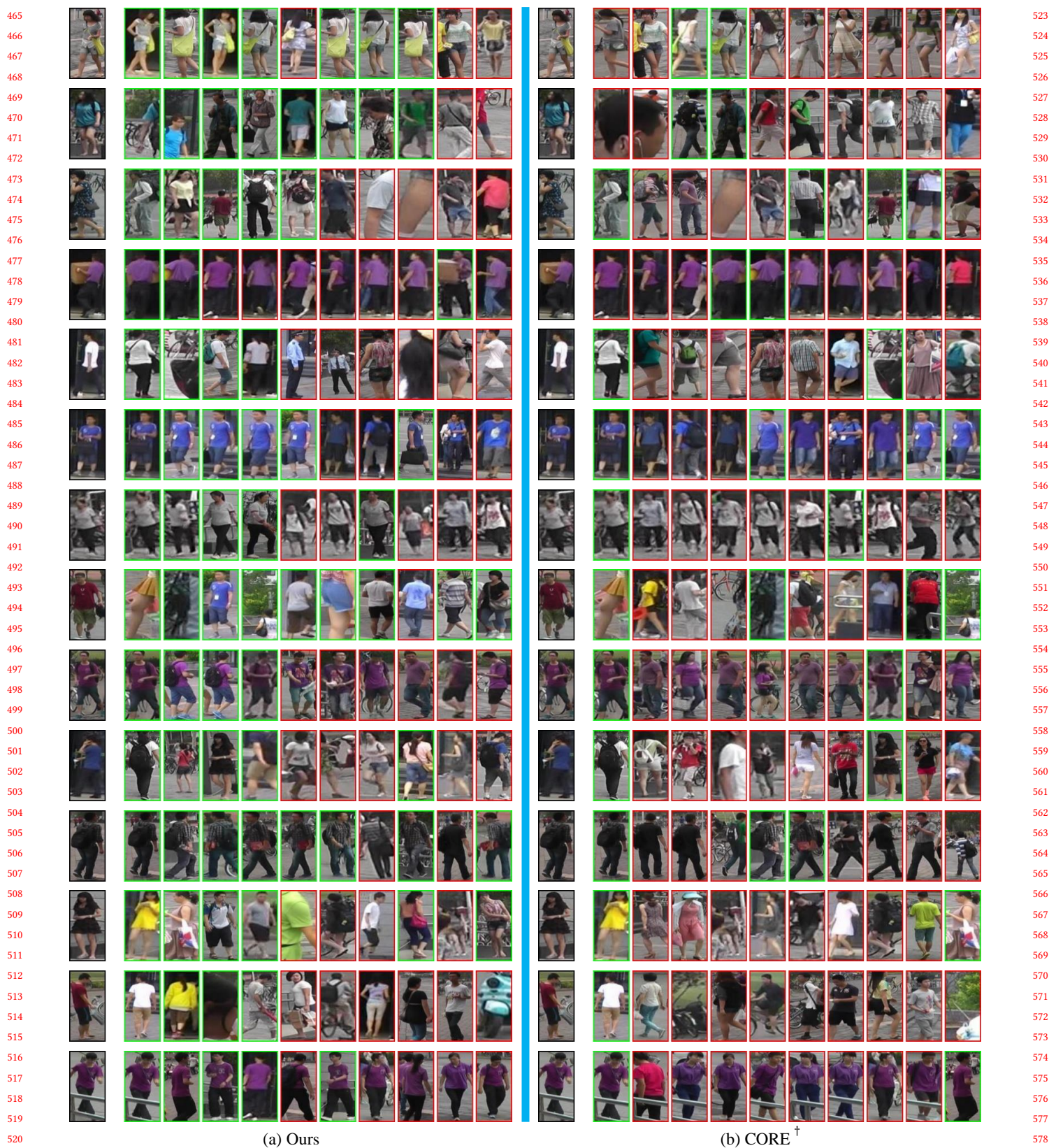


Figure 1: Person retrieval results of on seen domain Market-1501.

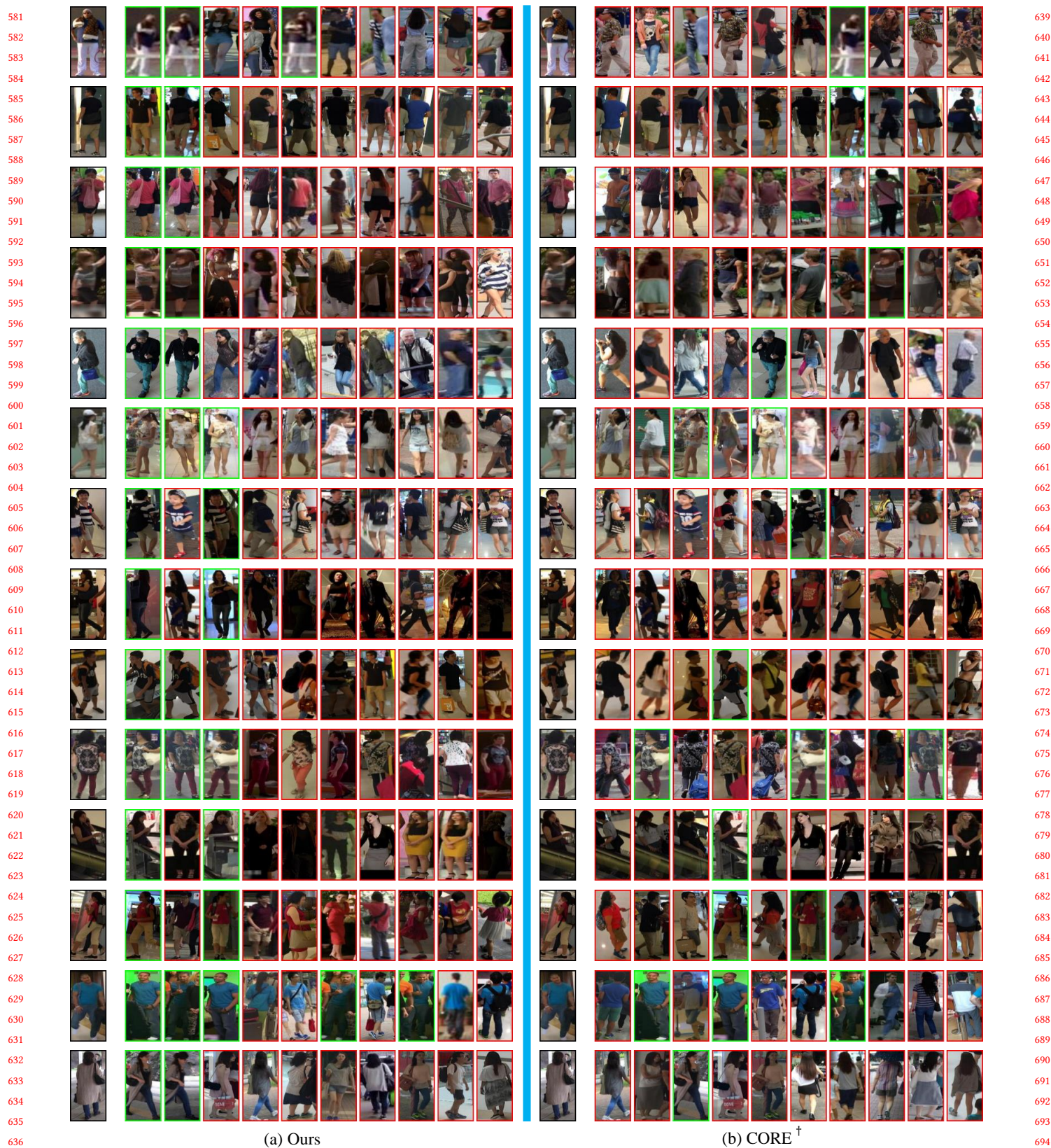


Figure 2: Person retrieval results of seen domain CUHK-SYSU.



Figure 3: Person retrieval results of seen domain MSMT17-V2.



Figure 4: Person retrieval results of seen domain CHUK03.



Figure 5: Person retrieval results of unseen domain CHUK01.

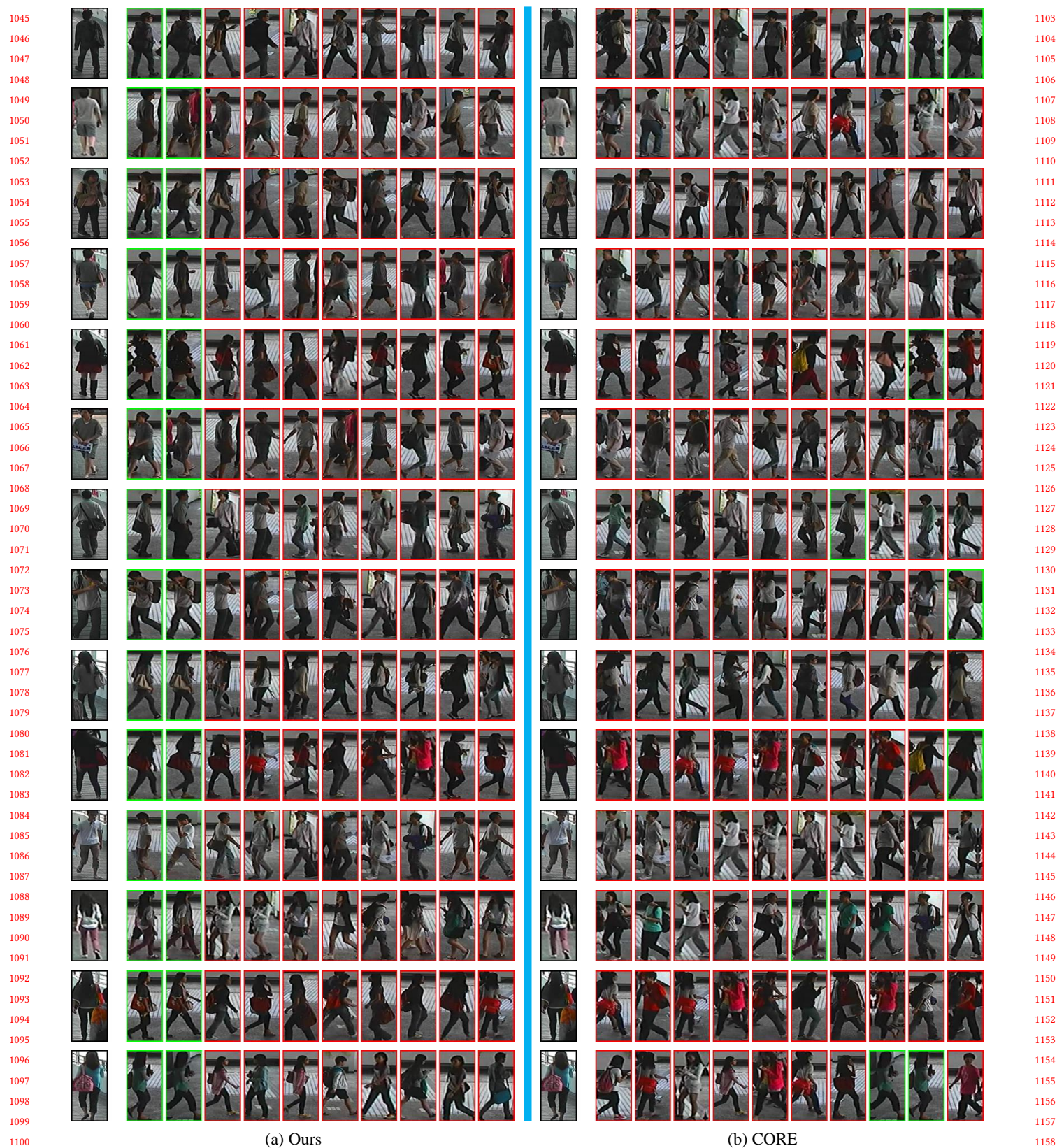


Figure 6: Person retrieval results of unseen domain CHUK02.

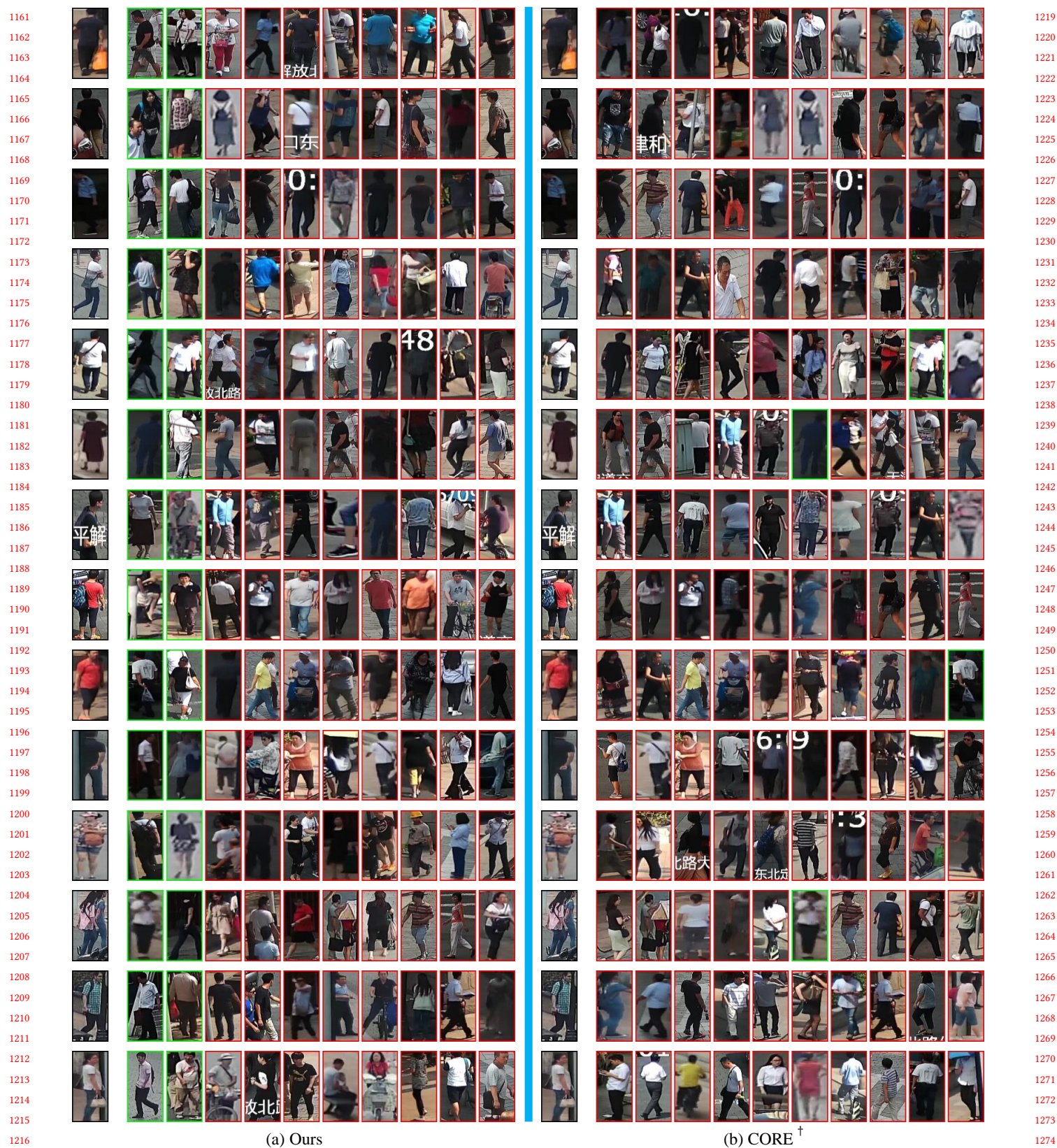


Figure 7: Person retrieval results of unseen domain SenseReID.

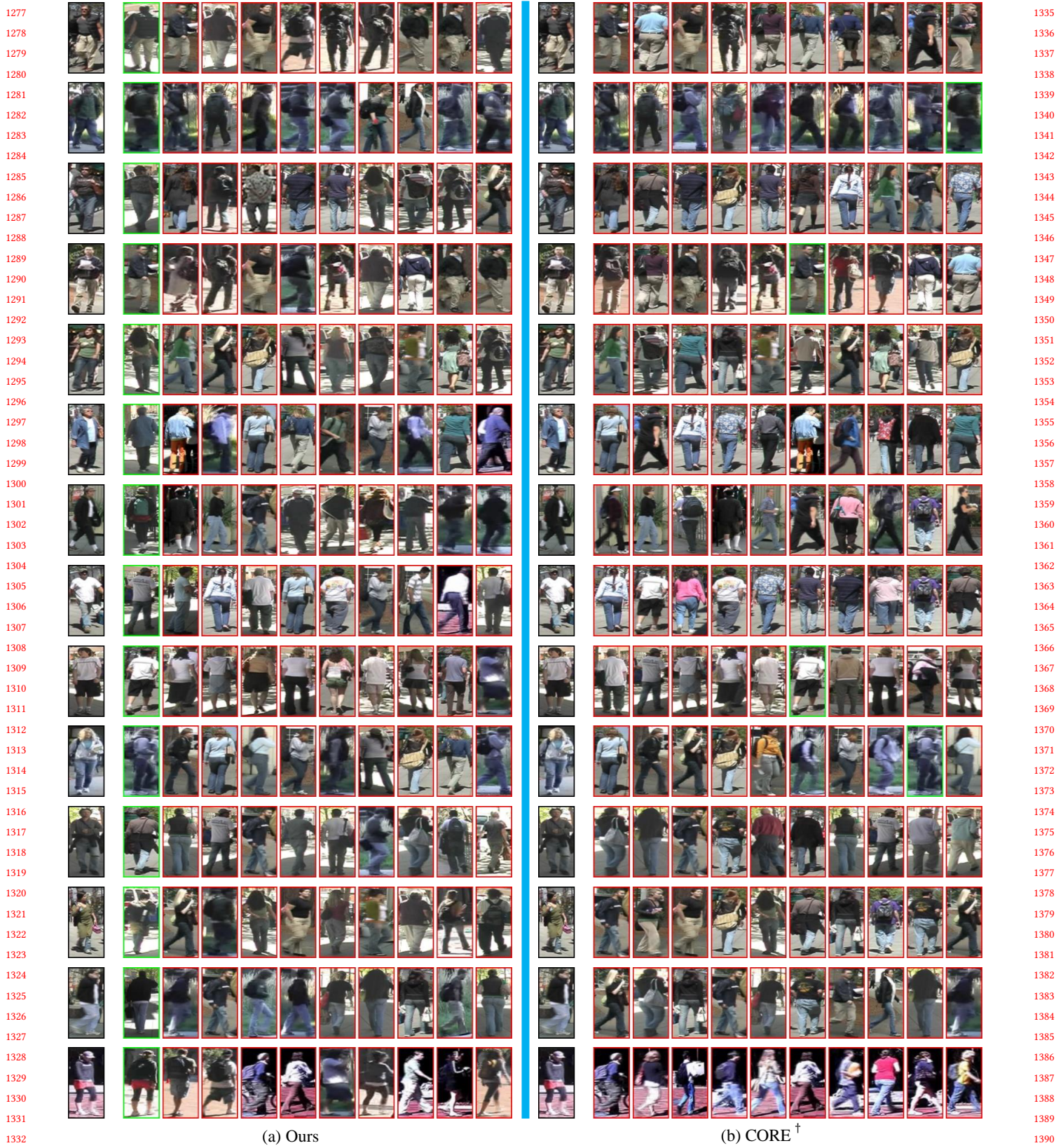


Figure 8: Person retrieval results of unseen domain VIPR.