Lifelong Person Re-Identification with Backward-Compatibility

Minyoung Oh[1] and Jae-Young Sim[1,2,†]
[1]Graduate School of Artificial Intelligence, [2]Department of Electrical Engineering,
Ulsan National Institute of Science and Technology, Republic of Korea
E-mail: {mmyy2513, jysim}@unist.ac.kr

***Abstract*—Lifelong person re-identification (LReID) assumes a practical scenario where the model is sequentially trained on continuously incoming datasets while alleviating the catastrophic forgetting in the old datasets. In such a case, not only the training datasets but the gallery images are also incrementally accumulated, that requires a huge amount storage space to store the gallery images as well as computational complexity to extract the features at the inference phase. In this paper, we address the above mentioned problem by incorporating the backward-compatibility to LReID based on the replay scheme. We attempt to train the model using the continuously incoming datasets while maintaining the model's compatibility toward the previously trained old models without re-computing the features of the old gallery images. Specifically, we design the cross-model compatibility loss based on the contrastive learning with respect to the replay features across all the old datasets. Experimental results demonstrate that the proposed method achieves a significantly higher performance of the backward-compatibility compared with the existing LReID methods.**

I. Introduction

Person re-identification (ReID) aims to retrieve specific individuals, corresponding to a given query, from the galleries of person images. It has drawn much attention in various practical applications such as surveillance and security. Note that real-world scenarios of ReID are subject to changing datasets of different domains which are collected at different places and time instances. Most of the conventional ReID methods, developed in the supervised learning manner, perform model retraining whenever a new dataset is given [1], [2]. Therefore, retraining the model by using the whole accumulated datasets together requires a huge amount of memory space and dramatically increases the computational complexity. To reduce the burden of entire retraining while maintaining the performance across different domains, the ReID models should learn the knowledge of the changing datasets incrementally.

For this purpose, lifelong person ReID (LReID) has been introduced [3]–[7], which sequentially trains the model on continuously incoming datasets across different domains. Specifically, it attempts to alleviate the catastrophic forgetting, where the model prefers the knowledge learned from the recent dataset while forgetting the old knowledge learned from the previously trained datasets. However, in practice, the challenges lie not only on the training but also on the inference. Whenever the model is updated with a new training dataset based on the LReID scenario, the feature space is also changed. Thus, at the inference stage, the updated model should re-compute the features from all the accumulated gallery images again to ensure that the features are compatible with that of the query images extracted by the new model. We call this process as *backfilling*. Since the gallery images are also continuously collected, the backfilling process takes a substantial amount of time especially for long-term deployment of ReID systems.

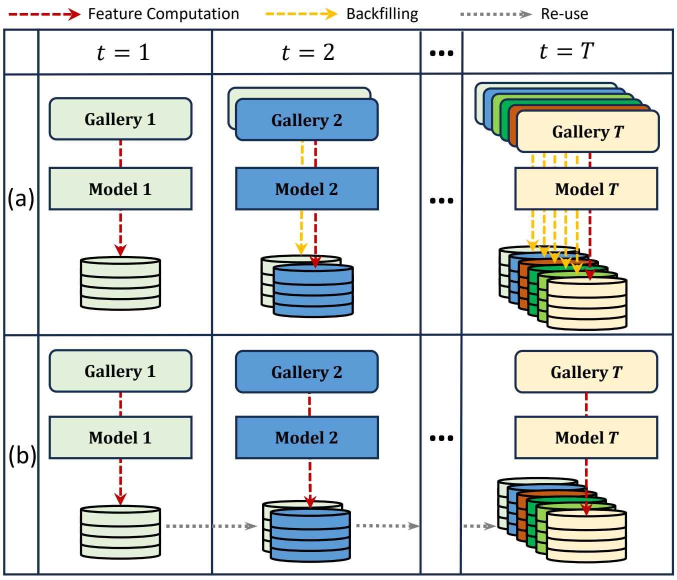

Fig. 1. Gallery feature computation at the inference phase. (a) The conventional LReID. (b) The proposed LReID with the backward-compatibility.

On the other hand, the backward-compatible training (BCT) has been investigated to address the time-consuming task of backfilling [8]–[10]. The BCT methods train the model by forcing the features computed by the new model to be close to the features obtained by the previously trained old models. Hence, based on BCT, the gallery features extracted by the old model can be re-used for re-identification with the new query features obtained by the new model, alleviating the burden of backfilling. However, the existing BCT methods jointly train all the accumulated datasets to update the model and do not consider the domain shift between the training datasets.

Based on the investigation of the complementary relationship between LReID and BCT, we devise a novel framework for LReID which ensures the backward-compatibility based on the replay dataset. Fig 1 illustrates the difference between (a) the conventional LReID and (b) the proposed LReID with the backward-compatibility in terms of the gallery feature computation at the inference phase. Whereas the conventional LReID conducts the backfilling process repeatedly whenever the model is updated with new datasets, the proposed framework considering the backward-compatibility uses the stored old gallery features, and hence is free from the time-consuming backfilling process. To ensure that the model has compatibility with the old models, we utilize the replay datasets and the corresponding features to design a cross-model compatibility loss inspired by the supervised contrastive loss. Experimental results show that the proposed method yields much better performance of LReID compared with the existing methods while achieving the backward-compatibility for inference.

Our main contributions are summarized as follows:

- We introduced a novel framework of LReID based on the replay dataset that avoids the time-consuming backfilling process and ensures the model's backward-compatibility at the inference stage.
- We proposed the cross-model compatibility loss by utilizing the replay images and the corresponding features.
- Experimental results showed that the proposed method provides better performance of LReID compared with the existing methods in terms of the backward-compatibility.

II. Related Works

A. Person Re-Identification

Most of the existing ReID methods have been developed in the supervised learning manner, that attempt to obtain ID-discriminative representation dealing with small inter-class variations and large intra-class variations. Sun et al. [11] horizontally divided the person image to extract part-aware features. Zeng et al. [12] learned the illumination-invariant features by disentangling the illumination and identity features. Luo et al. [13] enhanced the benefits of using the cross-entropy loss and the triplet loss. He et al. [14] introduced a transformer-based network leveraging the token-wise knowledge. Zhang et al. [1] employed a transformer-based network to augment the high-frequency patches and obtain fine-grained features. Ren et al. [15] utilized the attention mechanism to deal with the occlusion problem.

Recently, LReID has been studied to overcome the retraining burden of the supervised methods assuming more practical scenarios where the training datasets in different domains are continuously collected and provided. Pu et al. [4] adaptively accumulated and tranferred the knowledge by using the learnable knowledge graph. Wu and Gong [5] mitigated the catastrophic forgetting by ensuring knowledge coherence between the old and new models in terms of the classification, feature distribution and representation. Huang et al. [16] proposed an unsupervised domain-adaptation method for LReID that uses a meta-learning scheme to strike a balance between the new and old knowledge. Ge et al. [3] considered the LReID as a domain-adaptation problem where the old datasets are regarded as the source domain and the new dataset as the target domain, respectively. Yu et al. [7] made the old model trainable to refresh the old knowledge, which helps the model to consolidate the knowledge between the old and new models. Xu et al. [6] developed a replay-free LReID method by harmonically leveraging the long and short-term knowledge. However, the existing LReID methods consider the computational burden and memory consumption at the training stage only, while overlooking the burden of iterative feature computation on the accumulated gallery images at the inference stage.

B. Backward-Compatible Training

BCT was introduced to alleviate the backfilling burden of retrieval and search systems, such as the face recognition, whenever the model is updated with newly collected training datasets. The purpose of BCT is to train the new model to be compatible with the old models in terms of the feature extraction. Shen et al. [9] first introduced the concept of BCT where the features obtained by the new model are forced to be compatible with the frozen old classifier. Meng et al [17] proposed the transformation module which maps the features of one model into the other model's feature spaces by using the alignment and boundary losses. Wu et al. [10] enhanced the backward-compatibility by using the contrastive loss with respect to the features and logits of the new and old models. Pan et al. [8] performed the adversarial learning which makes the features, obtained by using different models, indistinguishable. Recently, Wan et al. [18] maintained the backward consistency in the retrieval networks based on the continual learning. However, it does not consider the changing datasets with huge distribution-shift across different domains, and the performance is only evaluated on the classes seen in the training which lacks the applicability to practical scenarios.

III. Proposed Method

Fig. 2 illustrates the training and inference procedures of the proposed LReID framework with backward-compatibility. Following the conventional LReID formulation with the replay scheme, we assume that T different training datasets are given in order as $\mathcal{D}_i\,(i = 1, 2, \cdots, T)$, and the model ϕ is sequentially trained by using the training datasets, respectively. Specifically, we first train the model ϕ_1 by using $\mathcal{D}_1$, and sample the replay data $\mathcal{R}_1$ from $\mathcal{D}_1$. We compute the features of $\mathcal{R}_1$ by using ϕ_1 and store them into $\boldsymbol{\mathcal{F}}_{\mathrm{R},1}$, which is then used to train the next models. At the T-th time instance, we train the model ϕ_T by using $\mathcal{D}_T$ and the previously sampled replay data $\{\mathcal{R}_i\}_{i=1}^{T-1}$. The pre-computed replay features $\{\boldsymbol{\mathcal{F}}_{\mathrm{R},i}\}_{i=1}^{T-1}$ are also used for backward compatibility training. Then we compute the features of the newly sampled replay data $\mathcal{R}_T$ by using the new model ϕ_T and store them into $\boldsymbol{\mathcal{F}}_{\mathrm{R},T}$.

At the inference stage, we test the model on the galleries, $\mathcal{G}_i\,(i = 1, 2, \cdots, T)$ with respect to the query set $\mathcal{Q}$. In the conventional LReID scenario, the features of all the accumulated galleries are computed again for inference whenever the model is newly updated. On the contrary, we store the features of the gallery $\mathcal{G}_i$ computed by using ϕ_i to $\boldsymbol{\mathcal{F}}_{\mathrm{G},i}$ which is used for inference at the next time instances. Consequently, after the training of the last (T-th) model, we use the new model ϕ_T

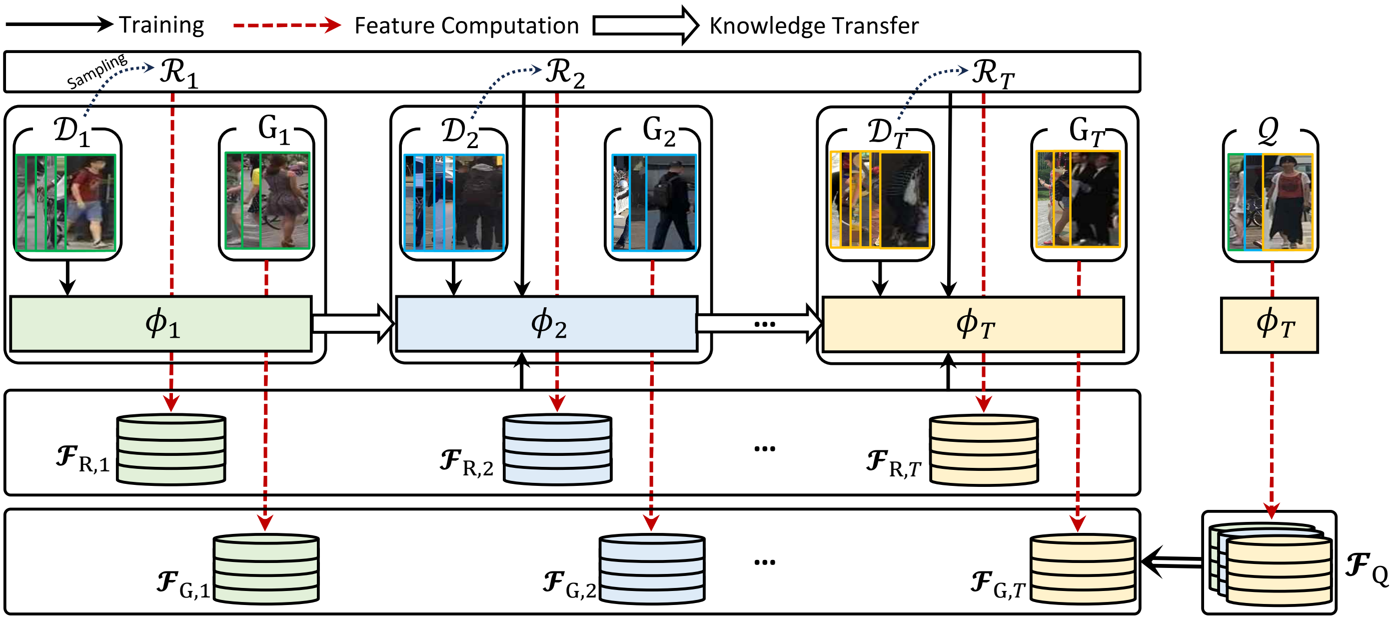

Fig. 2. Overview of the training and inference procedures in the proposed framework. We represent the feature extractors only for simplicity.

to compute the features of the gallery $\mathcal{G}_T$ only, and evaluate the performance of LReID on $\boldsymbol{\mathcal{F}}_{\mathrm{G},T}$ and the set of all the pre-computed gallery features $\{\boldsymbol{\mathcal{F}}_{\mathrm{G},i}\}_{i=1}^{T-1}$ together. Since we do not need to re-compute the features of previous galleries, it is a more practical setup without requiring the time-consuming backfilling process.

Whenever the model is sequentially trained on new datasets in the conventional LReID framework, the resulting feature spaces are changed to exhibit unique distributions according to the training domains. In such cases, the last model ϕ_T becomes incompatible with the previous models $\{\phi_1, \phi_2, ..., \phi_{T-1}\}$, and we should extract the features of all the galleries again by using ϕ_T. To ensure the compatibility of the new model with the previously trained models, we design the cross-model compatibility loss $\mathcal{L}_{\mathrm{cmc}}$ inspired by the concept of the supervised contrastive loss [19], given by

$$\mathcal{L}_{\mathrm{cmc}} = -\frac{1}{|B|}\sum_{k\in B}\log\left(\frac{\sum_{i=1}^{T-1}\sum_{j=1}^{|\boldsymbol{\mathcal{F}}^{k+}_{\mathrm{R},i}|}\exp\left\{\frac{\phi_T(\boldsymbol{r}_k)\cdot \boldsymbol{f}^{k+}_{i,j}}{\tau}\right\}}{\sum_{i=1}^{T-1}\sum_{j=1}^{|\boldsymbol{\mathcal{F}}_{\mathrm{R},i}|}\exp\left\{\frac{\phi_T(\boldsymbol{r}_k)\cdot \boldsymbol{f}_{i,j}}{\tau}\right\}+\mathcal{N}}\right), \tag{1}$$

where $\mathcal{N}$ is formulated as

$$\mathcal{N} = \sum_{j\in B_T}\exp\left\{\frac{\phi_T(\boldsymbol{r}_k)\cdot\phi_T(\boldsymbol{x}_j)}{\tau}\right\}. \tag{2}$$

B and B_T denote the mini-batches sampled from all the accumulated replay images and from the T-th dataset, respectively. $\boldsymbol{r}_k$ is the k-th image in B, and $\boldsymbol{x}_j$ is the j-th image in B_T. $\boldsymbol{\mathcal{F}}^{k+}_{\mathrm{R},i}$ is the subset of the replay features in $\boldsymbol{\mathcal{F}}_{\mathrm{R},i}$ positive to the identity of $\boldsymbol{r}_k$. $\boldsymbol{f}_{i,j}$ and $\boldsymbol{f}^{k+}_{i,j}$ denote the j-th replay features in $\boldsymbol{\mathcal{F}}_{\mathrm{R},i}$ and $\boldsymbol{\mathcal{F}}^{k+}_{\mathrm{R},i}$, respectively. τ is a temperature parameter. Note that the conventional supervised contrastive loss [19] takes the positive and negative samples within a mini-batch only without considering the model's backward-compatibility. On the contrary, we exploit the replay features across all the old datasets with different domains via $\{\boldsymbol{\mathcal{F}}_{\mathrm{R},i}\}_{i=1}^{T-1}$ based on the lifelong learning framework, and hence preserve the backward-compatibility more reliably by using the cross-model compatibility loss $\mathcal{L}_{\mathrm{cmc}}$ in (1).

Furthermore, the negative samples with respect to the each identity can help the model to improve the discriminative capability via the contrastive learning [21]. Based on the assumption that different datasets do not include the same identity simultaneously, we regard the images in the T-th dataset are the negative samples to the replay images, and minimize the similarity between the features of the old replay image $\boldsymbol{r}_k$ and the features of the new training images via the additional negative term in (2).

There has been an attempt [10] to achieve the backward-compatibility by using the supervised contrastive loss, which uses $\mathcal{R}$ and the old model ϕ_{T-1} instead of $\boldsymbol{\mathcal{F}}_{\mathrm{R}}$ where $\phi_{T-1}(\boldsymbol{r}_i)$ is utilized to guide the new model to represent the feature space obtained by ϕ_{T-1}. However, in the proposed lifelong learning formulation, the training datasets are dynamically changing across long duration of time instances, and thus it is hard for ϕ_{T-1} to keep the knowledge of the feature space at very old time instances suffering from the backward-compatibility of ϕ_T towards the old models. It is well-known that the LReID performance on the old datasets gets worse as T increases with more and more incoming datasets due to the catastrophic forgetting [22]. Note that the features of $\boldsymbol{\mathcal{F}}_{\mathrm{R}}$, extracted by the old models, preserve the representation of the old domains and thus guide the new model ϕ_T to ensure the backward-compatibility.

TABLE I
QUANITATIVE PERFORMANCE COMPARISON. WE RE-IMPLEMENTED THE EXISTING METHODS TO EVALUATE THE PERFORMANCE OF THE BACKWARD-COMPATIBILITY.

Train Order: MA → DU → CU → MS										
Method	MA		DU		CU		MS		Average	
	mAP	R-1	mAP	R-1	mAP	R-1	mAP	R-1	mAP	R-1
Joint-train	81.1	93.4	71.1	86.0	91.8	93.2	41.1	69.8	71.3	85.6
Fine-tuning	59.3	80.3	54.2	74.7	82.3	84.7	38.6	67.6	58.6	76.8
LwF [20]	62.9	82.0	56.8	74.8	78.8	81.2	13.5	33.7	53.0	67.9
BCT [9]	75.1	89.7	60.2	78.0	84.6	86.3	29.8	56.6	62.4	77.7
CVS [18]	70.7	87.0	56.3	72.0	81.5	83.8	27.2	53.0	58.9	74.0
PTKP [3]	73.4	86.4	**62.7**	77.8	83.8	85.3	36.3	62.0	64.1	77.9
KRKC [7]	59.5	80.3	49.3	69.7	80.7	82.4	36.7	65.3	56.6	74.4
LSTKC [6]	48.0	61.9	49.3	63.9	77.5	78.3	**42.7**	**68.7**	54.4	68.2
Proposed	**78.4**	**91.3**	62.5	**78.0**	**86.8**	**88.6**	33.1	60.4	**65.2**	**79.5**

IV. EXPERIMENTAL RESULTS

A. Experimental Setup

We used four benchmark datasets with unique domains for training: Market1501[23] (MA), DukeMTMC[24] (DU), CUHK-SYSU[25] (CU), and MSMT17[26] (MS), which are widely used in the existing literatures of LReID [3], [5], [27]. We pre-process the scene images in CUHK-SYSU for the purpose of ReID by using the ground-truth annotations following [5]. To evaluate the performance, we adopt the mean average precision (mAP) score and the Rank-1 (R-1) accuracy.

To train the network, we construct our baseline with $\mathcal{L}_{\text{base}}$ that is formulated as

$$\mathcal{L}_{\text{base}} = \lambda_{\text{ce}}\mathcal{L}_{\text{ce}} + \lambda_{\text{tri}}\mathcal{L}_{\text{tri}}, \tag{3}$$

where $\mathcal{L}_{\text{ce}}$ and $\mathcal{L}_{\text{tri}}$ are the cross-entropy loss and the hard triplet loss [28] calculated by using the feature extractor ϕ, respectively. λ_{ce} and λ_{tri} are hyper-parameters. We adopted the ImageNet pre-trained ResNet50 [29] as the feature extractor. Our total loss function to train ϕ_t is formulated as

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{base}} + \lambda_{\text{cmc}}\mathcal{L}_{\text{cmc}}, \tag{4}$$

where λ_{cmc} is a hyper-parameter. Hyper-parameters are empirically set as follows: $\lambda_{\text{ce}} = 1$ and $\lambda_{\text{tri}} = 1$ in (3), $\tau = 0.1$ in (1) and (2), and $\lambda_{\text{cmc}} = 0.05$ in (4). We followed [3] for other training details such as the replay memory update strategy, learning rate, etc.

B. Performance Comparison

To compare the performance of LReID with backward-compatibility, we re-implemented several existing methods: LwF [20], BCT [9], CVS [18], PTKP [3], KRKC [7], and LSTKC [6]. Table I shows the results of the proposed method compared with that of the existing methods. Joint-train means that we trained the model on all the datasets together, and Fine-tuning means that the model is independently trained on each new dataset in order. Both methods are trained with $\mathcal{L}_{\text{base}}$ only. The evaluation protocol is as follows. Each set of gallery features in $\{\boldsymbol{\mathcal{F}}_{\text{G,t}}\}_{t=1}^{T-1}$ is pre-computed by using the corresponding model of each task $\{\phi_t\}_{t=1}^{T-1}$, respectively, and we compute all the query features $\{\boldsymbol{\mathcal{F}}_{\text{Q,t}}\}_{t=1}^{T}$ and the features of the T-th gallery $\boldsymbol{\mathcal{F}}_{\text{G},T}$ by using ϕ_T after training on the last dataset.

We observe that the performance of almost methods on MSMT17, the last dataset, degrades compared to that of Fine-tuning due to the stability-plasticity dilemma [30] that the bias on the old model's knowledge disrupts the adaptation capability of the new knowledge. For instance, LwF suffers from the biasing toward the old knowledge and yields poor performance on MSMT17. On the one hand, PTKP achieves noticeable performance although it does not explicitly consider the backward-compatibility in training. Since PTKP casts the LReID problem into the domain adaption problem regarding the old dataset as source dataset, the model maps the features from the new dataset into the feature space of the old model that is trained with the source dataset. Thus, it could enhance backward-compatibility of the new model.

On the other hand, KRKC also addresses the LReID problem but it trains the old model as well unlike the other methods that freeze the parameters of the old model while preserving its feature space. Hence it exhibits poor backward-compatibility while adapting the new knowledge with the impressive performance on MSMT17 compared to the other methods. As the old model is trained, the feature space of old model is gradually forgotten, and the old model becomes no longer suitable for guiding the new model to ensure the backward-compatibility in KRKC. It is noteworthy that the existing LReID methods, e.g., PTKP, can enhance the backward-compatibility of the model, however it does not always guarantee the backward-compatibility as in the case of KRKC.

V. CONCLUSION

In this paper, we proposed a replay-based framework of LReID that ensures the backward-compatibility. Whereas the conventional LReID methods re-compute the features of the old gallery images whenever new datasets are given, the proposed method alleviates such time-consuming backfilling process by training the new model to be compatible toward the old domains. To this end, we designed the cross-model compatibility loss that utilizes the features of the replay images to guide the new model to represent the old feature space. Experimental results showed that the proposed method improves the performance of backward-compatibility in LReID by a large margin compared to the existing LReID methods.

ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (RS-2024-00392536), and by the Institute of Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (RS-2020-II201336, Artificial Intelligence Graduate School Program(UNIST)).

REFERENCES

[1] G. Zhang, Y. Zhang, T. Zhang, B. Li, and S. Pu, "Pha: Patch-wise high-frequency augmentation for transformer-based person re-identification," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recog.*, 2023, pp. 14 133–14 142.

[2] H. Luo, Y. Gu, X. Liao, S. Lai, and W. Jiang, "Bag of tricks and a strong baseline for deep person re-identification," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recog. Worksh.*, 2019.

[3] W. Ge, J. Du, A. Wu, *et al.*, "Lifelong person re-identification by pseudo task knowledge preservation," in *Proc. AAAI Conf. Artif. Intell.*, 2022, pp. 688–696.

[4] N. Pu, W. Chen, Y. Liu, E. M. Bakker, and M. S. Lew, "Lifelong person re-identification via adaptive knowledge accumulation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recog.*, 2021, pp. 7901–7910.

[5] G. Wu and S. Gong, "Generalising without forgetting for lifelong person re-identification," in *Proc. AAAI Conf. Artif. Intell.*, 2021, pp. 2889–2897.

[6] K. Xu, X. Zou, and J. Zhou, "Lstkc: Long short-term knowledge consolidation for lifelong person re-identification," in *Proc. AAAI Conf. Artif. Intell.*, vol. 38, 2024, pp. 16 202–16 210.

[7] C. Yu, Y. Shi, Z. Liu, S. Gao, and J. Wang, "Lifelong person re-identification via knowledge refreshing and consolidation," in *Proc. AAAI Conf. Artif. Intell.*, 2023, pp. 3295–3303.

[8] T. Pan, F. Xu, X. Yang, *et al.*, "Boundary-aware backward-compatible representation via adversarial learning in image retrieval," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recog.*, 2023, pp. 15 201–15 210.

[9] Y. Shen, Y. Xiong, W. Xia, and S. Soatto, "Towards backward-compatible representation learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recog.*, 2020, pp. 6368–6377.

[10] S. Wu, L. Chen, Y. Lou, *et al.*, "Neighborhood consensus contrastive learning for backward-compatible representation," in *Proc. AAAI Conf. Artif. Intell.*, 2022, pp. 2722–2730.

[11] Y. Sun, L. Zheng, Y. Yang, Q. Tian, and S. Wang, "Beyond part models: Person retrieval with refined part pooling (and a strong convolutional baseline)," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 480–496.

[12] Z. Zeng, Z. Wang, Z. Wang, Y. Zheng, Y.-Y. Chuang, and S. Satoh, "Illumination-adaptive person re-identification," *IEEE Trans. Multimedia*, vol. 22, no. 12, pp. 3064–3074, 2020.

[13] H. Luo, W. Jiang, Y. Gu, *et al.*, "A strong baseline and batch normalization neck for deep person re-identification," *IEEE Trans. Multimedia*, vol. 22, no. 10, pp. 2597–2609, 2019.

[14] S. He, H. Luo, P. Wang, F. Wang, H. Li, and W. Jiang, "Transreid: Transformer-based object re-identification," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2021, pp. 15 013–15 022.

[15] X. Ren, D. Zhang, X. Bao, and Y. Zhang, "S^2-net: Semantic and saliency attention network for person re-identification," *IEEE Trans. Multimedia*, vol. 25, pp. 4387–4399, 2023.

[16] Z. Huang, Z. Zhang, C. Lan, *et al.*, "Lifelong unsupervised domain adaptive person re-identification with coordinated anti-forgetting and adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recog.*, 2022, pp. 14 288–14 297.

[17] Q. Meng, C. Zhang, X. Xu, and F. Zhou, "Learning compatible embeddings," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2021, pp. 9939–9948.

[18] T. S. Wan, J.-C. Chen, T.-Y. Wu, and C.-S. Chen, "Continual learning for visual search with backward consistent feature embedding," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recog.*, 2022, pp. 16 702–16 711.

[19] P. Khosla, P. Teterwak, C. Wang, *et al.*, in *Proc. Adv. Neural Inform. Process. Syst.*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., 2020, pp. 18 661–18 673.

[20] Z. Li and D. Hoiem, "Learning without forgetting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 12, pp. 2935–2947, 2017.

[21] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in *Proc. Int. Conf. Mach. Learn.*, 2020, pp. 1597–1607.

[22] J. Kirkpatrick, R. Pascanu, N. Rabinowitz, *et al.*, "Overcoming catastrophic forgetting in neural networks," *Proc. Nat. Acad. Sci.*, vol. 114, no. 13, pp. 3521–3526, 2017.

[23] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, and Q. Tian, "Scalable person re-identification: A benchmark," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 1116–1124.

[24] E. Ristani, F. Solera, R. Zou, R. Cucchiara, and C. Tomasi, "Performance measures and a data set for multi-target, multi-camera tracking," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 17–35.

[25] T. Xiao, S. Li, B. Wang, L. Lin, and X. Wang, "Joint detection and identification feature learning for person search," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recog.*, 2017, pp. 3415–3424.

[26] L. Wei, S. Zhang, W. Gao, and Q. Tian, "Person transfer gan to bridge domain gap for person re-identification," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recog.*, 2018, pp. 79–88.
[27] J. Huang, X. Yu, D. An, *et al.*, "Learning consistent region features for lifelong person re-identification," *Pattern Recog.*, vol. 144, p. 109 837, 2023.
[28] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recog.*, 2015, pp. 815–823.
[29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recog.*, 2016, pp. 770–778.
[30] M. Mermillod, A. Bugaiska, and P. Bonin, "The stability-plasticity dilemma: Investigating the continuum from catastrophic forgetting to age-limited learning effects," *Frontiers Psychol.*, vol. 4, p. 504, 2013.