A. The detail of less forgetting performance

In this section, we investigate the relationship and influence between *forgetting* and *generalizing*. Corresponding to Fig. 3, we show the performance on first train domain over the followed four training steps in detail. As show in Tab. 6, the less forgetting rate δ, the better ability of generalization. Intuitively, remembering the knowledge learned from more identities is beneficial for generalising on unseen domain, because the accumulated previous identities are able to serve as full reference for comparing with new identities.

B. Illustration of LRe-ID evaluation protocol in a certain training step

As shown in Fig. 6, u is average performance on unseen domains, which is used to measure the capacity of generalising on unseen domains. s is average performance on seen domains, which is used to measure the capacity of retrieving incremental seen domains. Brown and black arrows denote 2^{nd} and 5^{th} training step, respectively. Assuming current step $t = 2$, the model is trained only on training data of the 2^{nd} domain. In each step, LReID tests the model on 1) testing sets of seen domains for evaluating domain forgetting, 2) testing set of current domain for evaluating current model's ability and 3) unseen identities in unseen domains for evaluating domain generalization ability through whole training process.

C. Feature Visualization

We extract 2-dimension t-SNE of features in the test set of unseen domains for visualising feature space. Comparing Fig. 7 and 8, over all training processes, LwF method does not improve generalization ability on unseen domain continually. In contrast, AKA optimizes feature space at every training step. In the last training step, the features space of AKA shows larger inter-class margin obviously.

D. The statistics of inter-class and intra-class distance

We calculate the frequency of inter-class and intra-class distance. In Fig. 9 and 10, the margin of these two distributions indicates the model's ability to improve intra-class compactness and inter-class separability. Our AKA outperforms LwF obviously.

Table 6: The investigation between forgetting and generalizing. The forgetting rate is simply calculated by $\delta = |max - min|$.

Method	Metric	MA→SY→DU→MS→CU (Test on MA)							DU→MS→MA→SY→CU (Test on DU)						
		t=1	2	3	4	5	δ	$\bar{u}$	t=1	2	3	4	5	δ	$\bar{u}$
SFT	mAP	70.6±0.2	44.7±0.3	27.4±0.4	23.2±0.2	24.1±0.2	47.4	35.2	62.8±0.2	33.6±0.3	30.0±0.2	19.6±0.3	12.9±0.2	49.9	34.8
	R-1	87.3±0.3	67.9±0.3	54.7±0.2	49.2±0.3	48.5±0.4	38.8	31.1	78.1±0.4	54.3±0.3	49.7±0.4	35.8±0.4	25.8±0.3	52.3	30.7
SFT-T	mAP	69.4±0.4	47.1±0.3	29.7±0.3	28.4±0.4	25.8±0.3	43.6	37.1	62.0±0.3	34.1±0.2	32.6±0.4	21.0±0.4	16.3±0.3	45.7	36.5
	R-1	86.7±0.5	70.5±0.4	53.5±0.4	55.3±0.4	47.9±0.5	38.8	34.2	77.8±0.5	55.8±0.4	51.9±0.5	37.5±0.4	29.2±0.5	48.6	33.9
SPD	mAP	70.6±0.2	54.1±0.3	34.6±0.4	32.8±0.2	30.5±0.3	40.1	36.3	62.8±0.2	42.1±0.3	37.8±0.2	26.3±0.4	20.8±0.3	42.0	35.3
	R-1	87.3±0.3	74.9±0.2	61.4±0.4	60.3±0.5	50.7±0.4	36.6	32.9	78.1±0.4	58.3±0.4	54.4±0.3	50.7±0.3	40.0±0.3	38.1	32.1
LwF	mAP	70.6±0.2	60.3±0.2	52.5±0.3	48.8±0.2	47.1±0.3	23.5	38.3	62.8±0.2	47.4±0.3	45.3±0.2	42.8±0.2	26.1±0.3	36.7	37.9
	R-1	87.3±0.3	76.5±0.2	74.9±0.1	71.8±0.3	65.1±0.4	22.2	36.9	78.1±0.4	63.4±0.2	60.7±0.3	59.2±0.3	44.6±0.2	33.5	34.8
CRL	mAP	70.6±0.2	63.1 ±0.5	55.9±0.3	52.4±0.4	48.5±0.5	22.1	38.5	62.8±0.2	50.1±0.3	47.9±0.3	45.3±0.2	29.8±0.3	33.0	38.0
	R-1	87.3±0.3	81.5±0.4	77.1±0.3	73.9±0.4	66.6±0.3	20.7	36.7	78.1±0.4	64.9±0.3	62.3±0.4	61.5±0.4	47.2±0.4	30.9	35.1
CRL-T	mAP	69.4±0.4	61.2±0.4	55.5±0.3	51.3±0.4	49.2±0.5	20.2	39.6	62.0±0.3	50.8±0.3	48.2±0.4	46.4±0.3	30.9±0.3	31.1	38.3
	R-1	87.0±0.5	79.4±0.2	76.5±0.4	73.3±0.5	67.0±0.3	20.0	38.1	77.8±0.5	65.0±0.4	62.5±0.6	62.0±0.5	48.2±0.5	29.6	35.5
AKA	mAP	68.7±0.4	65.9±0.4	59.8±0.3	56.7±0.3	51.2±0.2	17.5	44.3	61.8±0.2	52.2±0.3	51.5±0.3	48.4±0.3	32.5±0.3	29.3	40.8
	R-1	86.7±0.5	84.0±0.3	80.8±0.3	78.6±0.4	72.0±0.3	14.7	40.4	77.6 ±0.3	67.7±0.4	66.9±0.4	63.7±0.3	49.7±0.4	27.9	37.2

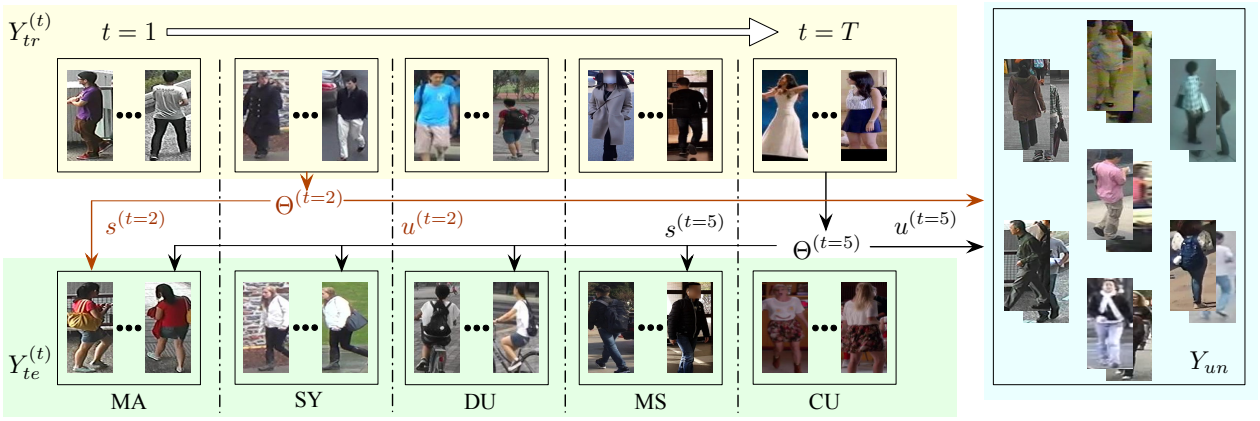
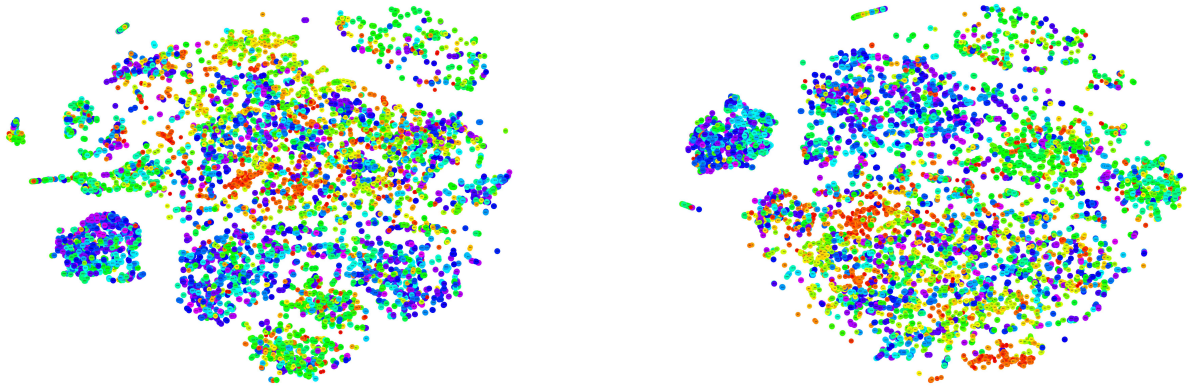


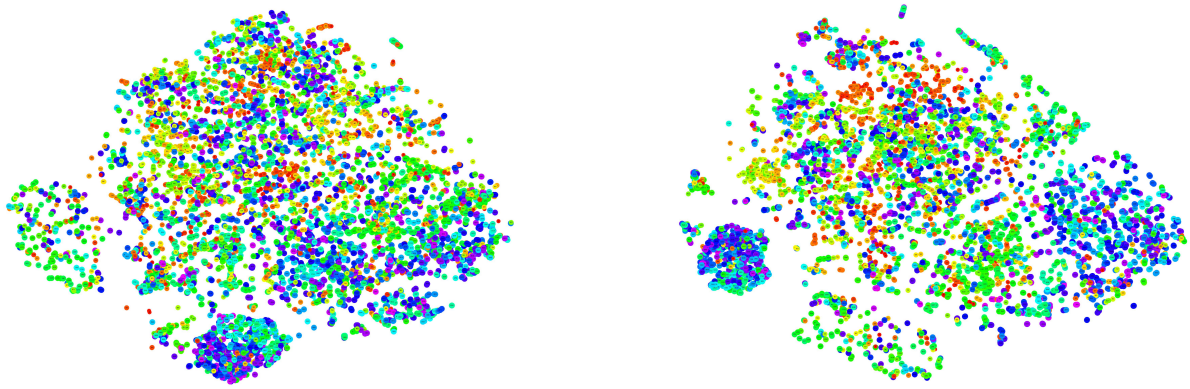
Figure 6: Illustration of LRe-ID evaluation protocol. u is average performance on unseen domains, which is used to measure the capacity of generalising on unseen domains. s is average performance on seen domains, which is used to measure the capacity of retrieving incremental seen domains. Brown and black arrows denote 2^{nd} and 5^{th} training step, respectively. Assuming current step $t = 2$, the model is trained only on training data of the 2^{nd} domain. In each step, LReID tests the model on 1) testing sets of seen domains for evaluating domain forgetting, 2) testing set of current domain for evaluating current model's ability and 3) unseen identities in unseen domains for evaluating domain generalization ability through whole training process.



(a) Result of LwF method in first training step.

(b) Result of LwF method in last training step.

Figure 7: Visualization of features in the test sets of unseen domains.



(a) Result of AKA method in first training step.

(b) Result of AKA method in last training step.

Figure 8: Visualization of features in the test sets of unseen domains.

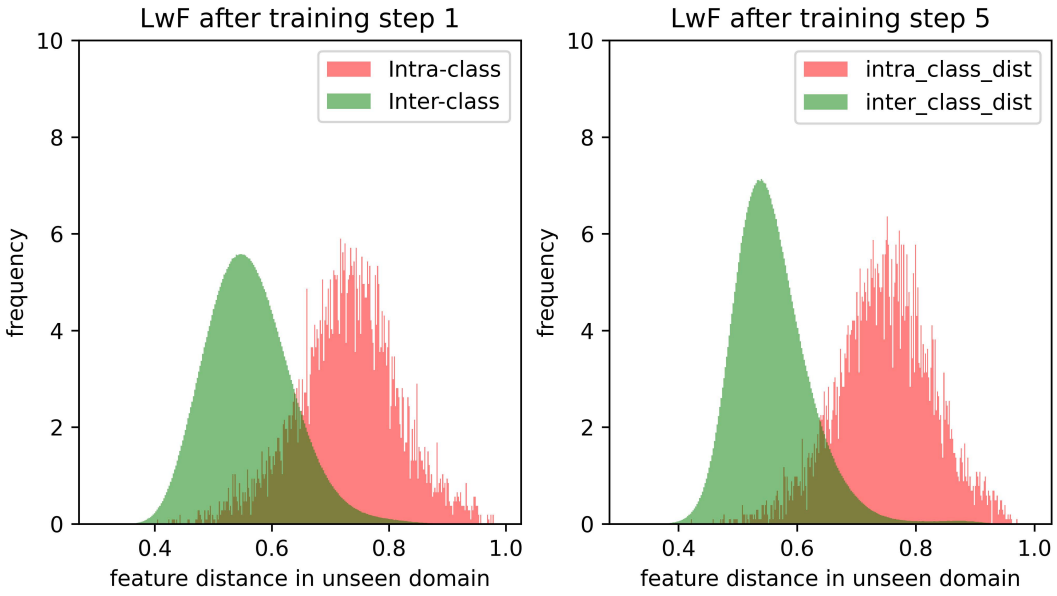


Figure 9: Histograms of inter-class and intra-class distance produced by LwF method.

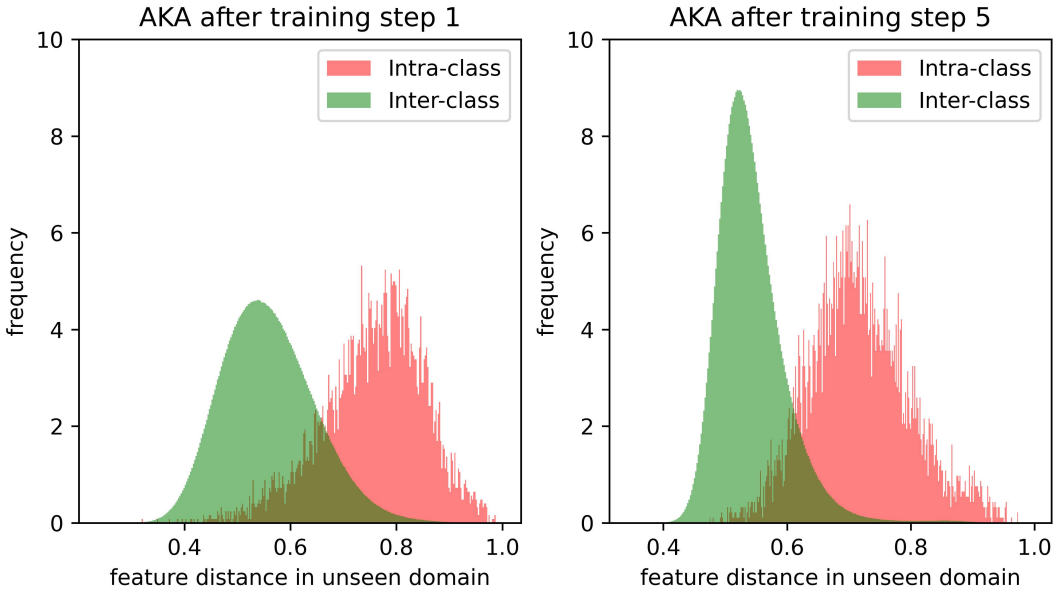


Figure 10: Histograms of inter-class and intra-class distance produced by our AKA method.