蚂蚁集团
ANT GROUP

ArXiv: https://arxiv.org/abs/2305.02610
Code: https://github.com/Ashespt/AdvBCT

Boundary-Aware Backward-Compatible Representation via Adversarial Learning in Image Retrieval (AdvBCT)

Tan Pan*, Furong Xu*, Xudong Yang, Sifeng He, Chen Jiang, Qingpei Guo,
Feng Qian, Xiaobo Zhang, Yuan Cheng, Lei Yang, Wei Chui

Poster Session: WED-PM-271

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

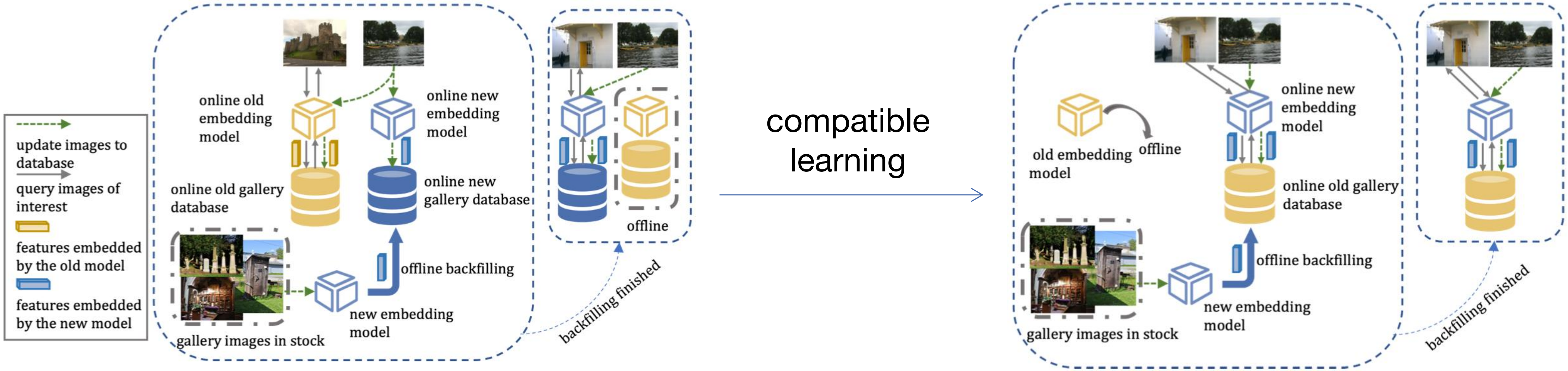
AdvBCT Overview:

- Goals: focus on the process of retrieval model upgrade.

1. Compatibility: Keep features consistent between two models to save time and resources during backfilling.
2. Disparity: Do not sacrifice the retrieval performance of the new model.

- Contributions:

1. Adversarial learning: minimize the gap between two feature spaces.
2. An elastic boundary constraint: balance disparity and compatibility.
3. Outperform other 4 methods in most allocation types .



(a) The process of the traditional upgrade.

(b) The process of the compatible upgrade.

Backward-Compatible Learning

Task: Ensure the compatibility of embedding representations between models. The new model can directly replace the old one and the embeddings of images in stock are updated on-the-fly.

Related Works:

1. point-to-point based(p2p):

constraint old embeddings and new embeddings

2. point-to-set based(p2s):

constraint old class centers and new embeddings

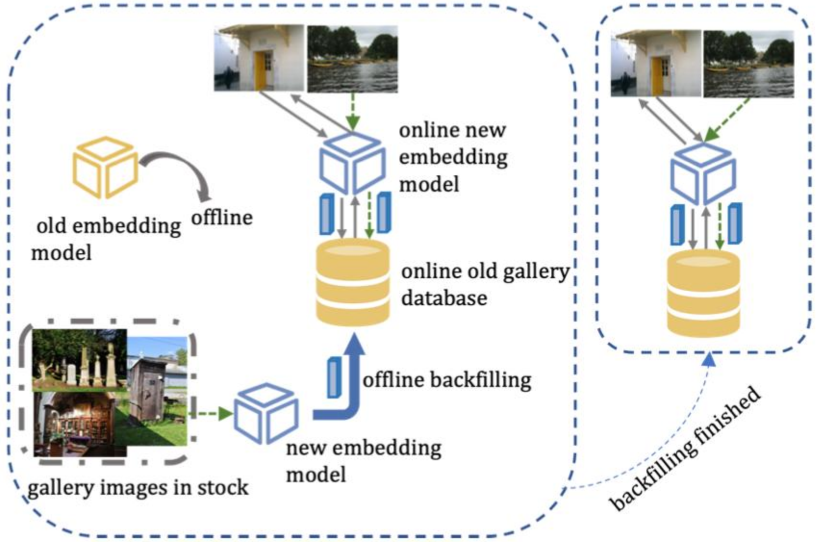


Figure 1. The process of the compatible upgrade

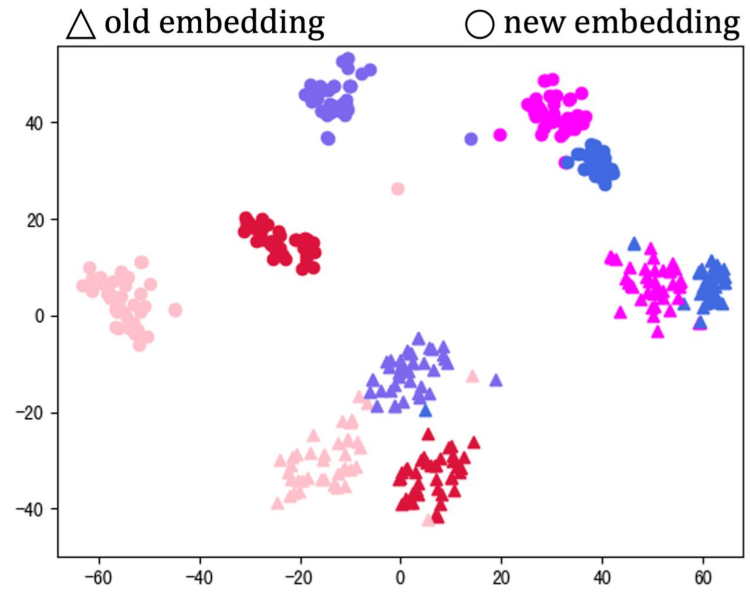
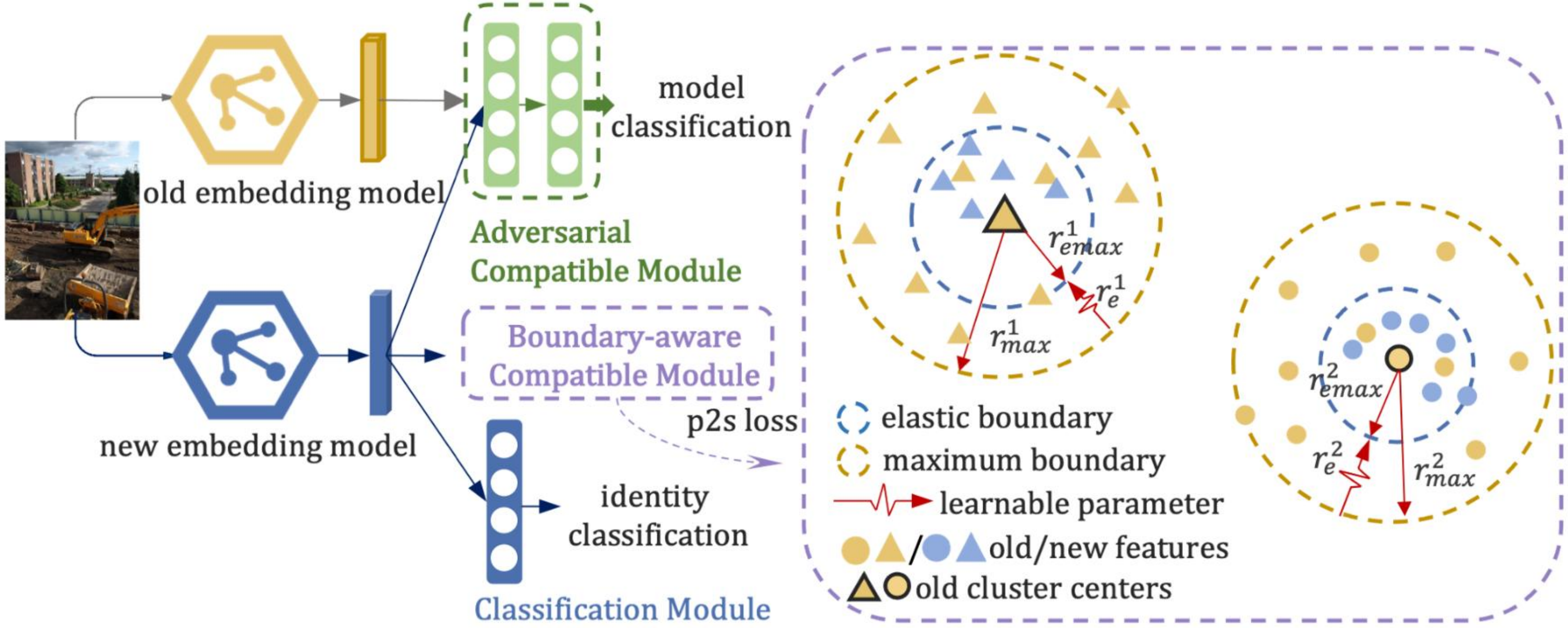


Figure 2. Distributions of the old embeddings and new embeddings without compatibility.

AdvBCT Pipeline:



Classification Module: Ensure the discrimination of new embeddings

$$L_{cls} = -\frac{1}{N}\sum_{i=1}^{N} y_i log p_i$$

Adversarial Compatible Module: A discriminator to minimize the gap between two embedding spaces of two models

$$\hat{\theta}_d = \underset{\theta_d}{\operatorname{argmin}} E(\hat{\theta}_n, \theta_d)$$

$$\hat{\theta}_n = \underset{\theta_n}{\operatorname{argmax}} E(\theta_n, \hat{\theta}_d)$$

Boundary-aware Compatible Module: Balance disparity and compatibility

Boundary-aware Compatible Module:

■ Why p2s constraint:

1. Applied to all pairs of samples, psp is influenced by outliers.
2. p2p constraint can be bounded by p2s constraint.

If compatible, the positive pairs and negative pairs should satisfy:

(p2p constraint)

$$\forall\{i,j,k\}, y_i = y_j \neq y_k$$

$$\langle\phi_n(x_i), \phi_o(x_j)\rangle < \langle\phi_n(x_i), \phi_o(x_k)\rangle$$

$$\langle\phi_n(x_i), \phi_o(x_j)\rangle < \langle\phi_n(x_i), \phi_n(x_k)\rangle$$

According to the triangle inequality:

(p2s constraint)

$$B_{lower} = \|\phi_n(x_i) - E_o(X^c)\|_2 - \|\phi_o(x_j) - E_o(X^c)\|_2 \quad (7)$$

$$B_{upper} = \|\phi_n(x_i) - E_o(X^c)\|_2 + \|\phi_o(x_j) - E_o(X^c)\|_2 \quad (8)$$

$$B_{lower} \leq \|\phi_n(x_i) - \phi_o(x_j)\|_2 \leq B_{upper} \quad (9)$$

■ Why elastic boundary:

Define $E_o(X)$ as cluster centers $\mathcal{O} = \{o^1, o^2, ..., o^C\}$ of the training data.

We can define p2s constraint as

$$\|\phi_n(x_i) - E_o(X^i)\|_2 < r^i_{max}$$

r^i_{max} represents the maximum distance of class i

Corner cases and outliers will make the distribution looser which leads to a larger $\boldsymbol{r^i_{max}}$.

$$r^k_e = w^k|r^k_{max} - t|$$

$$r^k_{emax} = \begin{cases} r^k_{max} - r^k_e & t < r_{max} \\ t - r^k_e & t > r_{max} \end{cases}$$

$$r^k_{emax} = \begin{cases} (1-w^k)r^k_{max} + w_k t & t < r_{max} \\ w^k r^k_{max} + (1-w_k)t & t > r_{max} \end{cases}$$

$$D_k = \sum_{i=1}^{m} max(\langle\phi_n(x^k_i), o^k\rangle - r^k_{emax}, 0)$$

$$L_{p2s} = \sum_{k=1}^{C} D_k$$

Benchmarks and Metrics

■ Data Allocation Types

- **Extended-data**:

$T_{old} = 30\%\ images, T_{new} = 100\%\ images$

- **Extenede-class**:

$T_{old} = 30\% classes, T_{new} = 100\% classes$

- **Extended-backbone (class)**:
  - Extended-class and $M_{old} = Res18$, $M_{new} = Res50$
- **Extended-backbone (data):**
  - Extended-data and $M_{old} = Res18$, $M_{new} = Res50$

Allocation type	Old train-set		New train-set	
	#images	#classes	#images	#classes
Extended-data	445,419	81,313	1,580,470	81,313
Extended-class	470,369	24,393	1,580,470	81,313
Extended-backbone (*class*)	445,419	81,313	1,580,470	81,313
Extended-backbone (*data*)	470,369	24,393	1,580,470	81,313

Train on GLDv2.
Test on RParis, ROxford, and GLDv2 test

■ Metrics

$$\mathcal{P}_{comp} = sigmoid(\frac{M(\phi_n, \phi_o; \mathcal{Q}, \mathcal{G}) - M(\phi_o, \phi_o; \mathcal{Q}, \mathcal{G})}{M(\phi_*, \phi_*; \mathcal{Q}, \mathcal{G}) - M(\phi_o, \phi_o; \mathcal{Q}, \mathcal{G})}) \quad (16)$$

$$\mathcal{P}_{up} = sigmoid(\frac{M(\phi_n, \phi_n; \mathcal{Q}, \mathcal{G}) - M(\phi_*, \phi_*; \mathcal{Q}, \mathcal{G})}{M(\phi_*, \phi_*; \mathcal{Q}, \mathcal{G})}) \quad (17)$$

$$\mathcal{P}_{\beta-score} = \frac{(1+\beta^2)\mathcal{P}_{comp} * \mathcal{P}_{up}}{\beta^2 \mathcal{P}_{comp} + \mathcal{P}_{up}} \quad (18)$$

$M(\phi_o, \phi_n; \mathcal{Q}, \mathcal{G})$ represents the mAP with the setting that embeddings of $\mathcal{Q}$ and embeddings of G are extracted by ϕ_n and ϕ_o respectively.

Experiments

Allocation type	Model_{old}	Model_{new}	RParis self	RParis cross	ROxford self	ROxford cross	GLDv2-test self	GLDv2-test cross	$\mathcal{P}_{up}$	$\mathcal{P}_{comp}$	$\mathcal{P}_{1-score}$
Extended-data	ϕ_o^{R18}	-	75.45	-	49.15	-	10.03	-	-	-	-
	-	ϕ_*^{R18}	81.15	4.93	63.85	1.20	16.48	0.2	-	7.19	-
	ϕ_o^{R18}	ϕ_{BCT}^{R18}	80.58	77.37	56.34	49.66	14.61	11.30	48.02	54.71	51.13
	ϕ_o^{R18}	ϕ_{LCE}^{R18}	**81.57**	77.83	**60.85**	**51.35**	**16.48**	**12.17**	**49.65**	57.41	53.34
	ϕ_o^{R18}	ϕ_{UniBCT}^{R18}	80.93	**78.30**	57.24	50.97	**16.06**	**13.25**	48.90	**59.19**	**53.52**
	ϕ_o^{R18}	$\phi_{Hot-refresh}^{R18}$	79.57	76.53	58.15	50.05	13.88	10.35	47.78	52.50	50.03
	ϕ_o^{R18}	ϕ_{AdvBCT}^{R18}	**82.78**	**78.55**	**62.13**	**52.31**	15.71	11.49	**49.55**	**58.09**	**53.45**
Extended-class	ϕ_o^{R18}	-	74.29	-	54.34	-	11.43	-	-	-	-
	-	ϕ_*^{R18}	81.15	4.93	63.85	1.29	16.48	0.2	-	3.38	-
	ϕ_o^{R18}	ϕ_{BCT}^{R18}	79.45	76.13	58.94	53.43	14.79	**12.26**	48.33	52.79	50.41
	ϕ_o^{R18}	ϕ_{LCE}^{R18}	**81.26**	**76.78**	**60.49**	54.29	**16.07**	12.04	**49.37**	53.95	**51.51**
	ϕ_o^{R18}	ϕ_{UniBCT}^{R18}	76.92	74.55	59.07	**57.82**	14.80	**12.31**	48.09	**54.78**	51.17
	ϕ_o^{R18}	$\phi_{Hot-refresh}^{R18}$	78.93	75.33	60.31	51.68	14.0	10.41	48.06	47.26	47.57
	ϕ_o^{R18}	ϕ_{AdvBCT}^{R18}	**82.05**	**77.16**	**64.51**	**54.82**	**16.44**	12.05	**50.16**	**54.87**	**52.35**
Extended-backbone (*data*)	ϕ_o^{R18}	-	75.45	-	49.15	-	10.03	-	-	-	-
	-	ϕ_*^{R50}	87.66	4.81	76.56	2.26	22.12	0.2	-	15.44	-
	ϕ_o^{R18}	ϕ_{BCT}^{R50}	85.54	78.95	68.66	54.42	19.11	12.78	47.81	55.86	51.52
	ϕ_o^{R18}	ϕ_{LCE}^{R50}	**87.49**	**79.36**	**75.16**	57.18	**21.85**	**13.30**	**49.73**	**57.31**	**53.25**
	ϕ_o^{R18}	ϕ_{UniBCT}^{R50}	84.37	78.91	67.42	56.18	**21.57**	**13.42**	48.48	56.80	52.31
	ϕ_o^{R18}	$\phi_{Hot-refresh}^{R50}$	86.78	78.81	75.10	**57.84**	20.61	12.41	49.19	56.53	52.60
	ϕ_o^{R18}	ϕ_{AdvBCT}^{R50}	**86.79**	**79.08**	**75.71**	**59.03**	20.77	12.75	**49.31**	**57.30**	**53.01**
Extended-backbone (*class*)	ϕ_o^{R18}	-	74.29	-	54.34	-	11.43	-	-	-	-
	-	ϕ_*^{R50}	87.66	4.81	76.56	2.26	22.12	0.2	-	11.46	-
	ϕ_o^{R18}	ϕ_{BCT}^{R50}	84.18	77.12	68.93	57.52	18.85	13.47	47.71	54.59	50.91
	ϕ_o^{R18}	ϕ_{LCE}^{R50}	**85.77**	77.31	72.89	**58.95**	**20.84**	**14.00**	**49.04**	**55.67**	**52.15**
	ϕ_o^{R18}	ϕ_{UniBCT}^{R50}	82.48	**77.33**	66.37	58.51	18.91	**13.99**	47.29	55.51	51.06
	ϕ_o^{R18}	$\phi_{Hot-refresh}^{R50}$	85.62	77.29	**73.95**	56.83	20.47	12.44	49.01	53.63	51.21
	ϕ_o^{R18}	ϕ_{AdvBCT}^{R50}	**86.24**	**78.31**	**73.93**	**60.33**	**22.03**	13.54	**49.65**	**56.64**	**52.83**

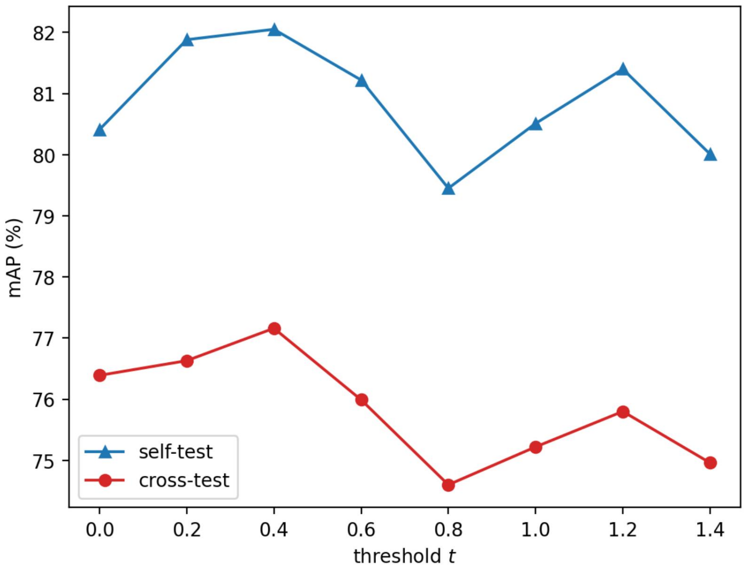
The compatible training benchmark testing on BCT, LCE, Hot-refresh, UniBCT, and AdvBCT

Experiments

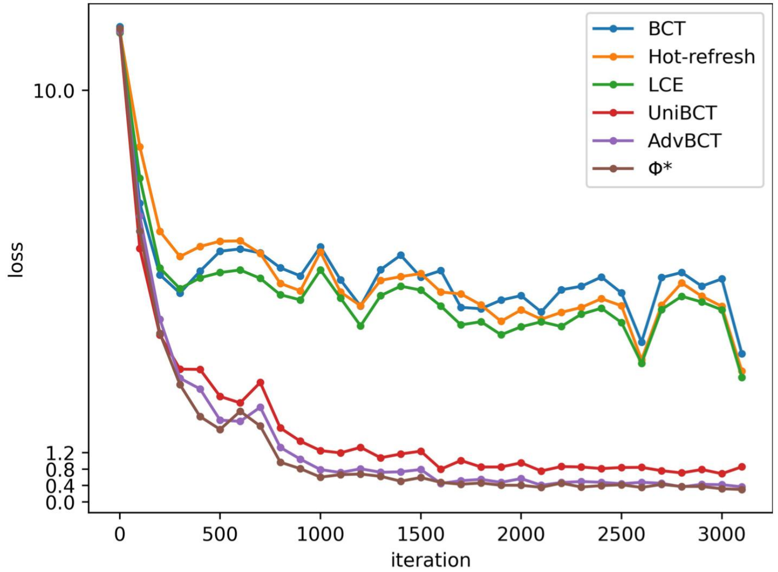
■ Ablation Study

#	$\mathcal{L}_{cls}$	$\mathcal{L}_{adv}$	$\mathcal{L}_{p2s}$	RParis self	RParis cross	ROxford self	ROxford cross	RParis self	RParis cross	ROxford self	ROxford cross
1(ϕ_o)	✓			75.45	-	49.15	-	74.29	-	54.34	-
2(ϕ_*)	✓			81.15	4.93	63.85	1.29	81.15	4.93	63.85	1.29
3		✓		5.8	6.69	3.25	1.66	5.14	6.8	2.44	1.76
4			✓	76.4	75.23	50.78	44.75	74.82	74.13	51.09	48.22
5	✓	✓		80.87	4.4	**63.92**	2.11	81.09	5.67	62.3	2.34
6	✓		✓	82.12	77.18	61.16	51.63	81.66	76.16	63.59	52.92
7		✓	✓	76.83	75.7	52.76	49.31	75.09	74.45	51.39	48.67
8	✓	✓	✓	**82.78**	**78.55**	62.13	**52.31**	**82.05**	**77.16**	**64.51**	**54.82**

Comparison results of different components in Extended-data (left) and Extended-class (right)



The influence of parameter t in Extended-class



The converge trend of 5 methods

Conclusions

1. To better ensure compatibility, we designed the adversarial and boundary-aware compatible modules.
2. Adversarial compatible module aims to pull the embedding distributions of the old and new models close.
3. Boundary-aware compatible module is used to obtain a suitable boundary to constrain distance relationship between the new and old embeddings
4. We establish a comprehensive benchmark for subsequent researchers to handily contribute to the field.

Thanks for watching