Domain-Aware Augmentations for Unsupervised Online General Continual Learning

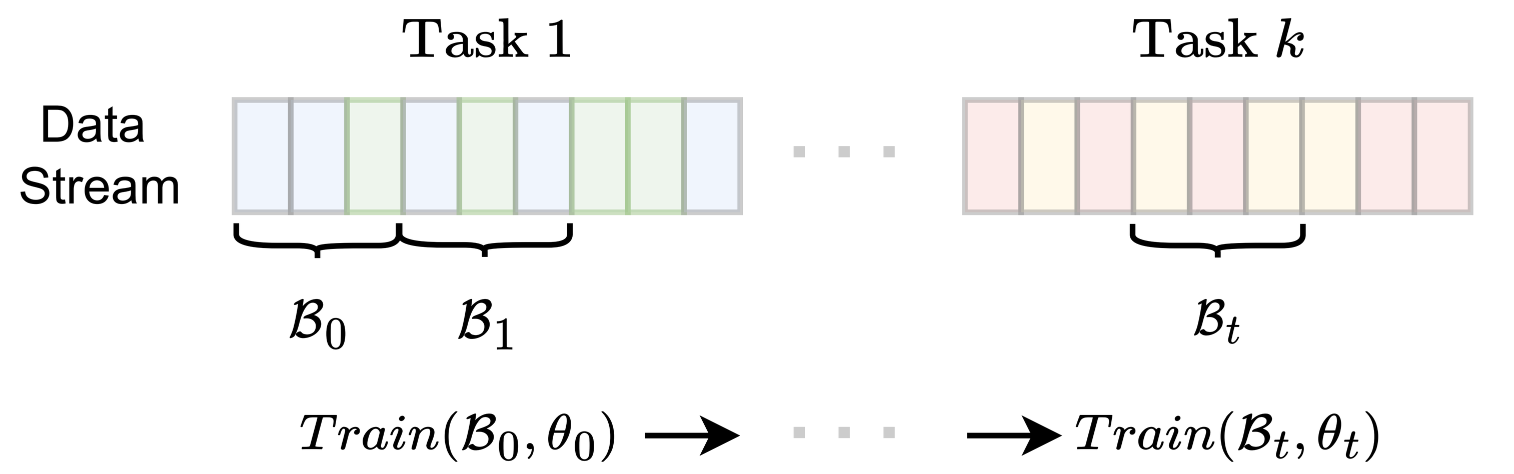
Nicolas Michel   Romain Negrel   Giovanni Chierchia   Jean-François Bercher

Univ Gustave Eiffel, CNRS, LIGM, ESIEE Paris

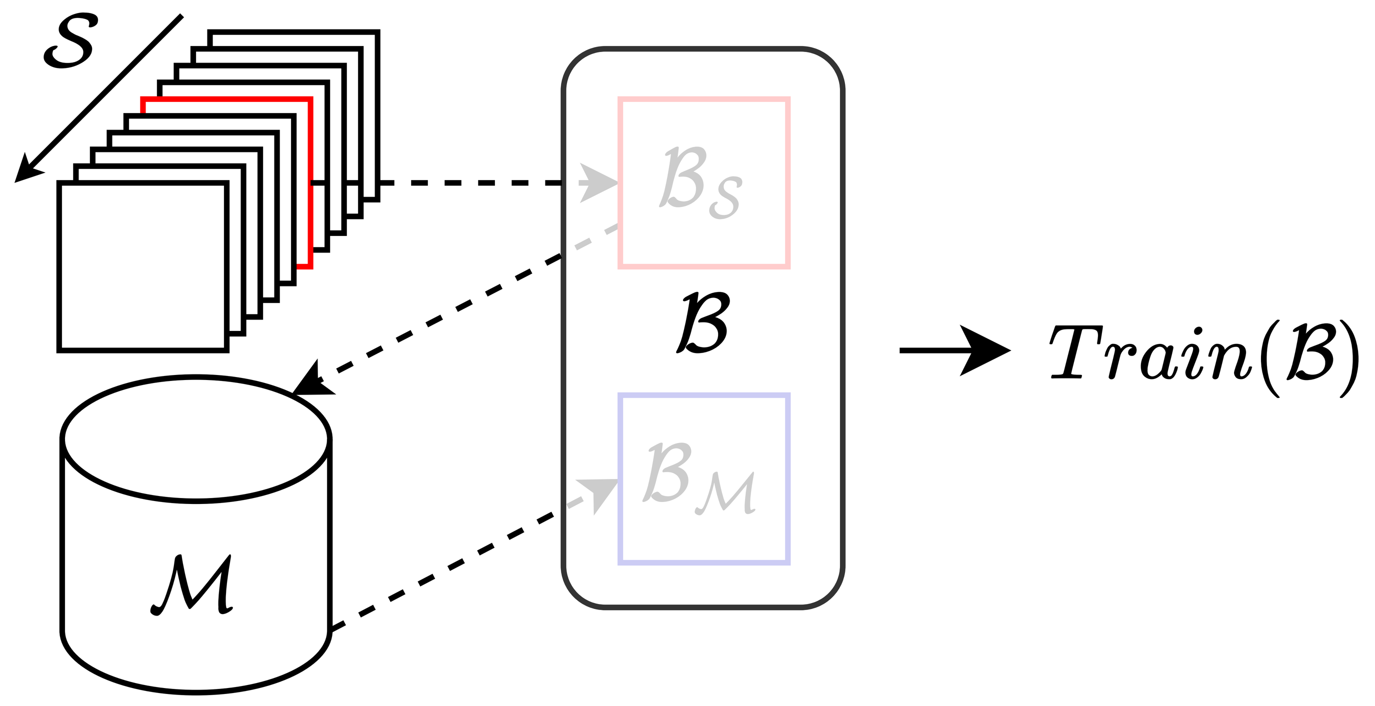
Problem Setting

Online Continual Learning

- K tasks: $\{\mathcal{T}_1, \dots, \mathcal{T}_K\}$ with K datasets $\{\mathcal{D}_1, \dots, \mathcal{D}_K\}$
- $\mathcal{D}_k = \{X_k\}$, unlabeled data
- Only one pass over the data is allowed
- Task change is unknown



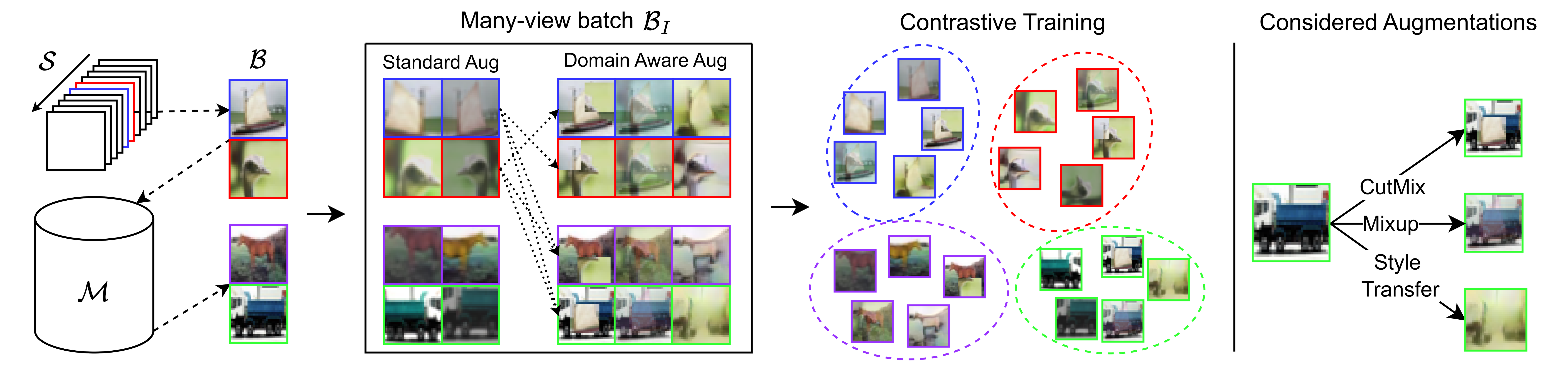
Memory-based methods



- Memory based approaches are state-of-the-art in OCL.

Description of the approach

Overview



Domain Aware Augmentations

Augmentation strategies

Consider $x_i \in \mathcal{M}$ from memory and $x_d \in \mathcal{S}$ from stream.

Domain-Aware Mixup:

$$x_a = \lambda \cdot x_i + (1 - \lambda) \cdot x_d$$

Domain-Aware CutMix:

$$x_a = M \odot x_i + (\mathbf{1} - M) \odot x_d$$

with M a binary mask constructed according to $B = (r_x, r_y, r_w, r_h)$, defined with $r_w = W\sqrt{1-\lambda}$, $r_h = H\sqrt{1-\lambda}$.

$\lambda \sim \mathcal{U}(0.5, 1)$ such that the augmented image x_a has a minimum amount of information coming from x_i

Experiments

Datasets & Baselines

CIFAR100, CIFAR10. Tiny ImageNet: 110 000 64x64 images; 200 classes.

Baselines

- Supervised: Experience replay (ER) Rolnick et al. 2019, Supervised Contrastive Replay (SCR) Mai et al. 2021, Gradient Self-Adaptation (GSA) Guo et al. 2023.
- Unsupervised: Self Taught Associative Memory (STAM) Smith 2019, Simple Siamese (SimSiam) Chen 2021, Bootstrap Your Own Latent (BYOL) Grill 2020, Simple Contrastive Learning of Representations (SimCLR) Chen 2020.

Results

		CIFAR10		CIFAR100		Tiny ImageNet	
	Method	$M=200$	$M=500$	$M=2k$	$M=5k$	$M=5k$	$M=10k$
Sup.	ER	41.46 ± 3.41	52.93 ± 4.39	31.37 ± 0.69	39.22 ± 1.11	19.4 ± 2.26	25.93 ± 3.02
	SCR	49.16 ± 3.02	60.28 ± 1.21	37.79 ± 0.95	47.31 ± 0.34	28.80 ± 0.51	34.28 ± 0.28
	GSA	52.03 ± 2.14	61.30 ± 2.35	38.77 ± 1.07	48.21 ± 0.99	27.58 ± 0.74	34.72 ± 0.82
Unsupervised	STAM	30.54 ± 0.8		8.39 ± 0.4		-	
	SimSiam-ER	27.73 ± 1.18	30.59 ± 1.21	6.91 ± 0.37	7.47 ± 0.11	6.49 ± 0.41	6.9 ± 0.52
	BYOL-ER	29.43 ± 0.55	29.30 ± 1.01	9.39 ± 0.52	10.35 ± 0.61	6.19 ± 0.26	6.59 ± 0.38
	SimCLR-ER	43.20 ± 2.30	48.81 ± 0.78	21.2 ± 0.9	23.62 ± 0.54	16.7 ± 0.5	17.97 ± 0.14
	Ours $(7,1,0,0,0)$	45.68 ± 2.38	52.89 ± 0.57	27.27 ± 0.13	31.32 ± 0.64	17.9 ± 0.58	20.21 ± 0.13
	Ours $(4,1,1,1,1)$	$\mathbf{48.09 \pm 1.22}$	$\mathbf{56.02 \pm 1.34}$	$\mathbf{29.02 \pm 0.77}$	$\mathbf{33.19 \pm 0.9}$	$\mathbf{20.35 \pm 0.02}$	$\mathbf{22.06 \pm 0.37}$

Table. Final Average Accuracy (%)

Conclusions

- Domain Aware Augmentations outperform standard augmentation strategies;
- Close performances to supervised baselines when memory size is small;
- Optimizing memory usage is key to improving memory-based approach in Online Continual Learning.