WACV 2025
TUCSON, ARIZONA • FEB 28 - MAR 4

Towards Unbiased Continual Learning: Avoiding Forgetting in the Presence of Spurious Correlations

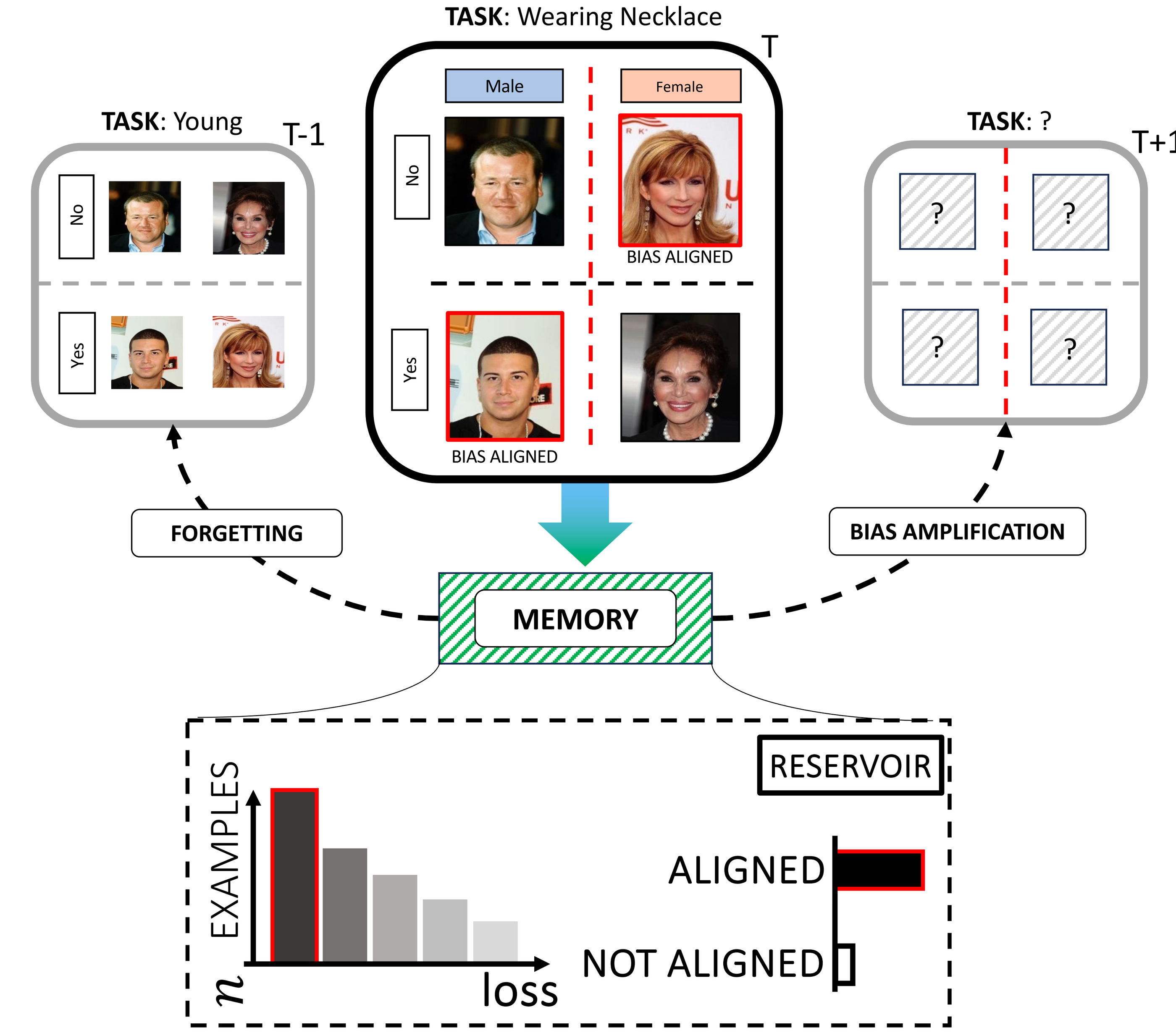
Giacomo Capitani, Lorenzo Bonicelli, Angelo Porrello,
Federico Bolelli, Simone Calderara, Elisa Ficarra

University of Modena and Reggio Emilia, Italy email: *name.surname*@unimore.it

AImageLab University of Modena and Reggio Emilia
UNIMORE UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA
MAMMOTH

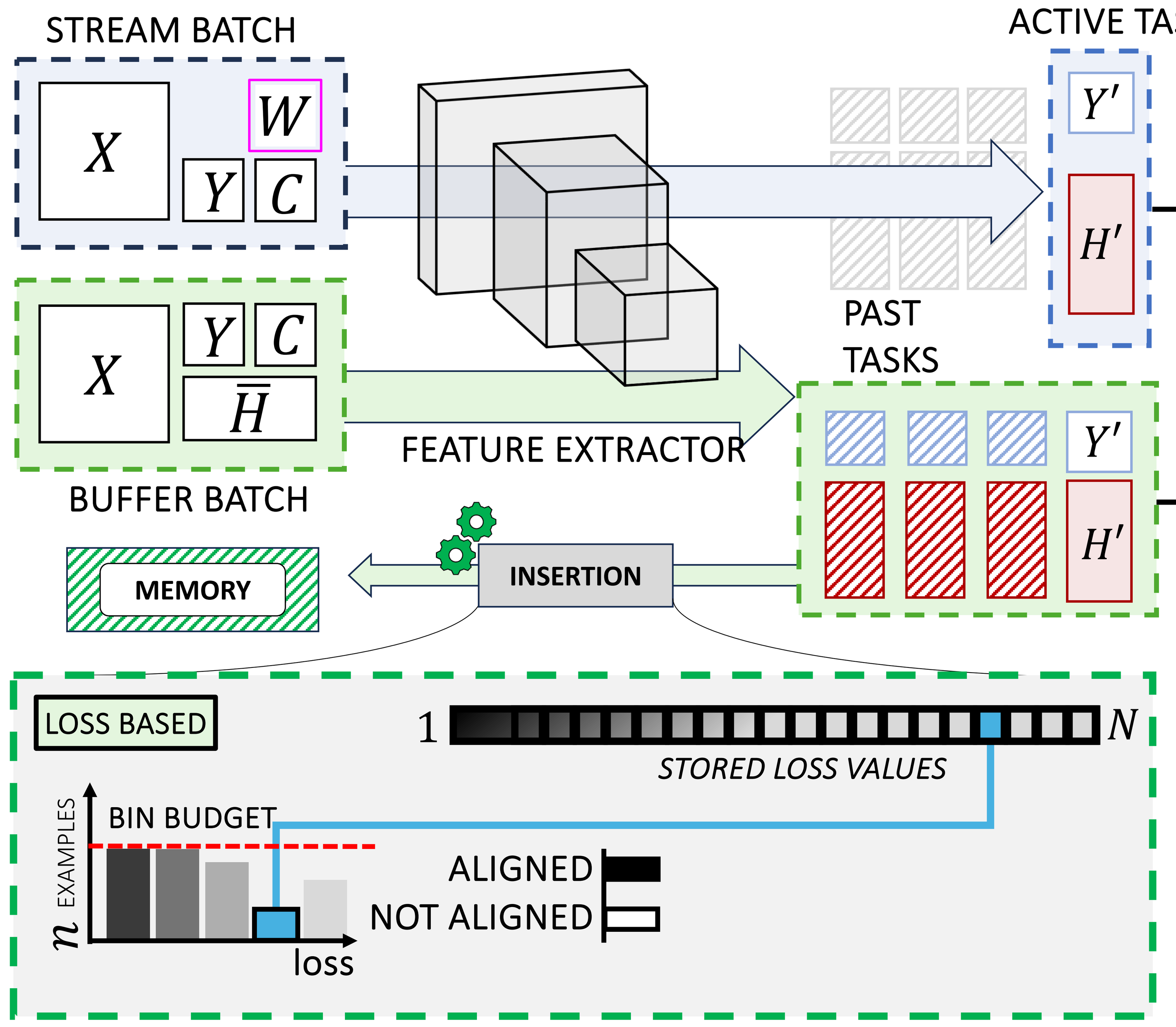
PROBLEM STATEMENT

- Rehearsal methods is largely used in **Continual Learning** (CL) to manage the insertion of new samples in the buffer memory and to overcome forgetting.
- As the buffer serves as the sole source of information on past tasks, a buffer filled with spurious correlations may amplify existing biases, creating a compounding effect.



! Biases can significantly impair the efficacy of **CL** models by inducing reliance on suboptimal shortcuts during data stream and memory retention, exacerbating catastrophic forgetting.

METHOD



Stream Objective

$$L_{cluster} = L_{CE}(C, H') \quad \text{Eq. 1)}$$

$$L_{target} = w_c L_{BCE}(Y, Y') \quad \text{Eq. 2)}$$

$$\boxed{w_c = \frac{1}{N} E_{(x,y)\sim P_c}\left[L_{BCE}(Y,Y') + \gamma L_{CE}(Y,Y')\right]} \quad \text{Eq. 3)}$$

Buffer Objective

$$L_{buf} = L_{BCE}(Y, Y') + L_{BCE}(C, H') \quad \text{Eq. 4)}$$

$$KD_{buf} = E_{(x,\overline{H})\sim M}\left[\,\|\overline{H} - H'\|_2^2\right] \quad \text{Eq. 5)}$$

Sample Re-weighting. For each cluster we compute w_c, which is proportional to its average classification error (Eq. 3).

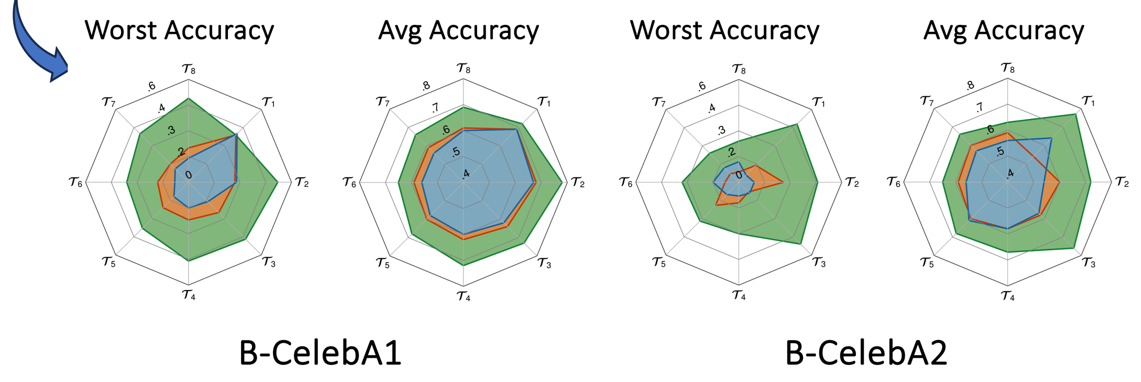
Buffer Insertion. Our loss-based insertion mechanism aims to ensure fair representation inside the memory

EXPERIMENTAL RESULTS

Method	B-CelebA1 Acc$_{worst}$[%]	B-CelebA1 Acc$_{avg}$[%]	B-CelebA2 Acc$_{worst}$[%]	B-CelebA2 Acc$_{avg}$[%]	B-Camelyon Acc$_{worst}$[%]	B-Camelyon Acc$_{avg}$[%]
Random	50.00	50.00	50.00	50.00	50.00	50.00
SGD	14.87 ± 1.56	60.12 ± 0.68	8.12 ± 0.57	56.06 ± 0.09	48.53 ± 6.47	85.8 ± 1.51
Debiasing						
BPA	15.08 ± 1.56	61.69 ± 0.47	9.16 ± 0.47	56.33 ± 0.53	62.13 ± 2.73	88.06 ± 1.11
CFIX	18.00 ± 2.04	64.00 ± 1.25	17.65 ± 1.97	61.26 ± 0.96	59.56 ± 0.83	87.88 ± 0.57
Replay (1024)						
BGS†	55.68 ± 2.92	74.64 ± 0.34	56.56 ± 2.74	76.45 ± 0.51	77.55 ± 0.07	91.89 ± 0.41
ER-ACE	16.37 ± 1.76	60.75 ± 0.77	13.12 ± 1.10	59.03 ± 0.59	56.80 ± 1.70	88.48 ± 0.08
DER++	21.79 ± 1.06	61.34 ± 0.54	18.03 ± 1.37	60.87 ± 0.36	53.40 ± 1 41	82.56 ± 0.57
BPA + replay	16.04 ± 0.90	60.92 ± 0.51	11.37 ± 0.43	58.19 ± 0.66	65.33 ± 1 02	88.88 ± 0.52
CFIX + replay	17.80 ± 0.04	61.57 ± 0.39	19.79 ± 0.75	62.62 ± 0.58	55.93 ± 0.34	86.48 ± 0.58
LwP	19.40 ± 0.91	62.33 ± 1.31	13.44 ± 3.94	57.47 ± 1.47	54.40 ± 9.54	86.39 ± 4.24
LwS (ours)	**58.84 ± 2.42**	**71.43 ± 1.67**	**50.91 ± 3.52**	**70.73 ± 0.65**	**80.40 ± 1.74**	**92.47 ± 0.21**

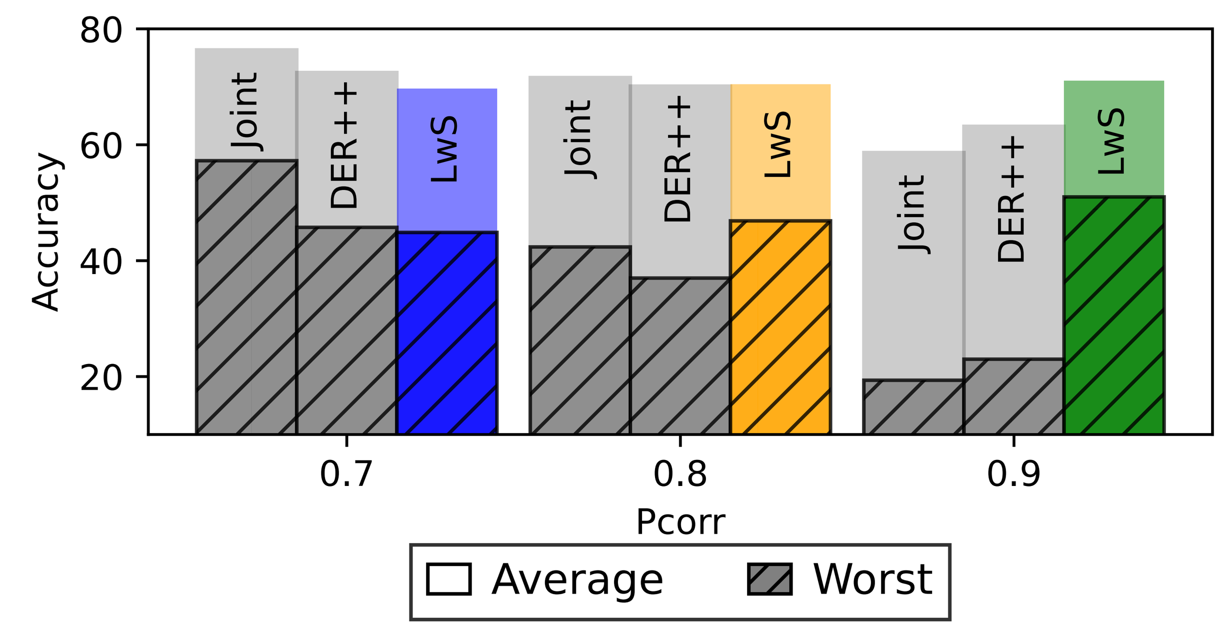
! **LwS** boosts average and worst-group accuracy out-performing standard rehearsal methods.

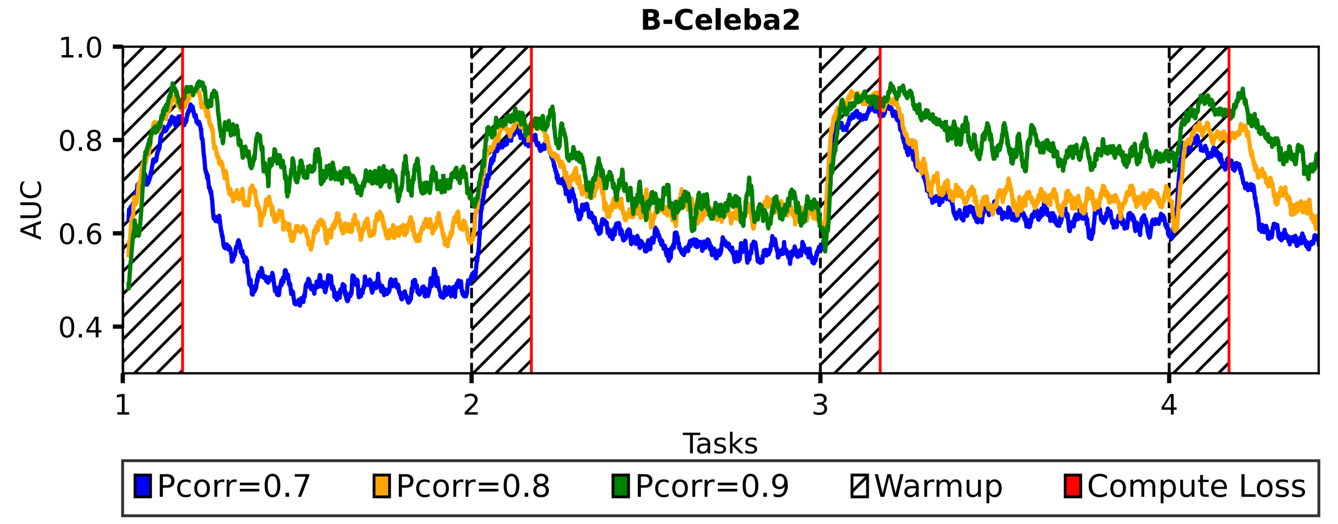
! **Learning Trend.** Comparative analysis across tasks, showcasing worst-group and avg accuracy.



ABLATIONS

- **Increasing the magnitude of spurious correlations (p_corr).** As spurious correlations increases, standard methods (even the Joint training) drop in average and worst-case accuracy while our approach remain stable.





- **Fixing $w_c = 1$ (left)** worsened model performance. **LwS using Resevoir vs Loss-based (right).** Adopting the loss-based approach consistently increased robustness, in terms of worst-group and average accuracy.

Dataset	w_c	Acc$_{worst}$	Acc$_{avg}$
B-CelebA1	adaptive	**58.84 ± 2.42**	**71.43 ± 1.67**
	fixed	52.83 ± 1.59	70.90 ± 0.99
B-CelebA2	adaptive	**50.91 ± 3.52**	70.73 ± 0.65
	fixed	47.29 ± 1.43	**70.74 ± 0.68**
B-Camelyon	adaptive	**80.40 ± 1.74**	**92.47 ± 0.21**
	fixed	78.20 ± 1.60	91.85 ± 0.37

	$\mathcal{M}$	Strategy	Acc$_{worst}$[%]	Acc$_{avg}$[%]
B-CelebA1	256	reservoir	14.14	58.21
		loss-based	**36.29**	**66.73**
	512	reservoir	18.50	61.08
		loss-based	**52.12**	**71.17**
	1024	reservoir	17.87	62.16
		loss-based	**56.98**	**72.57**