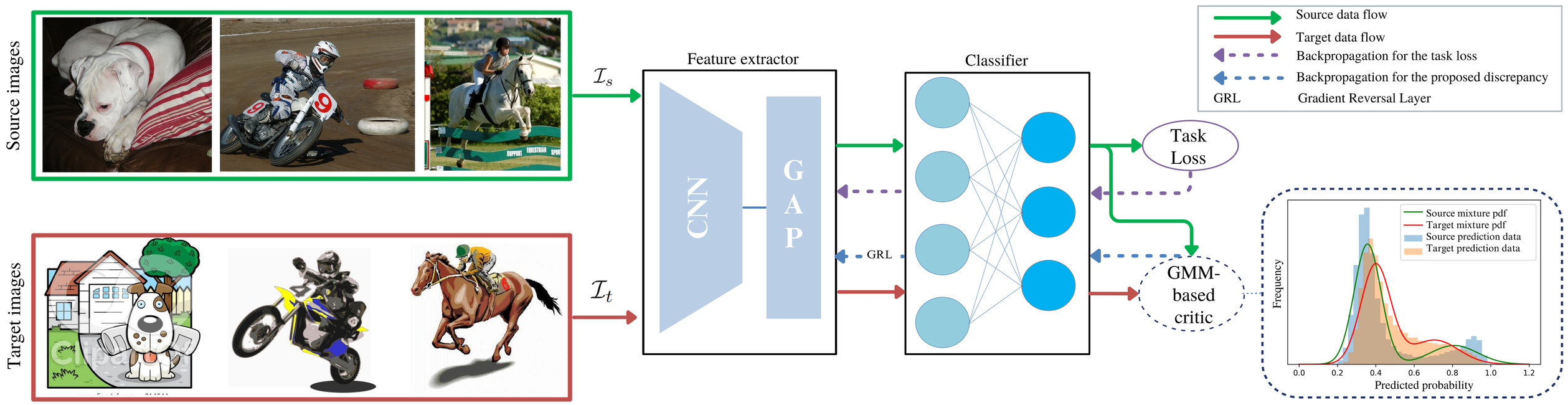
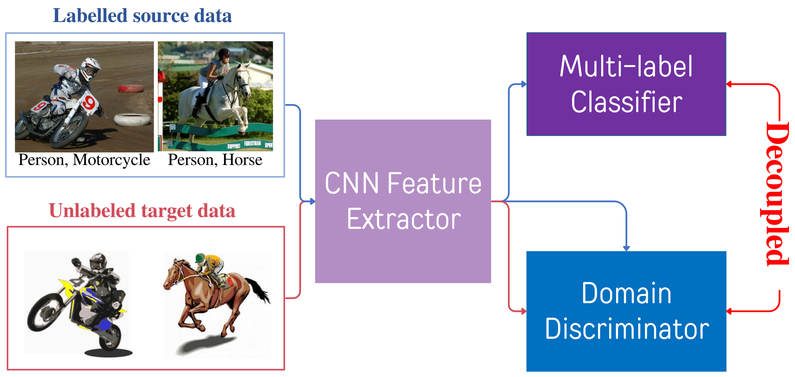
Discriminator-free Unsupervised Domain Adaptation for Multi-label Image Classification

Inder Pal Singh[1] inder.singh@uni.lu
Enjie Ghorbel[1,2] enjie.ghorbel@uni.lu
Anis Kacem[1] anis.kacem@uni.lu
Arunkumar Rathinam[1] arunkumar.rathinam@uni.lu
Djamila Aouada[1] djamila.aouada@uni.lu

[1]Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg
[2]High Institute of Multimedia Arts (ISAMM), University of Manouba, Tunisia

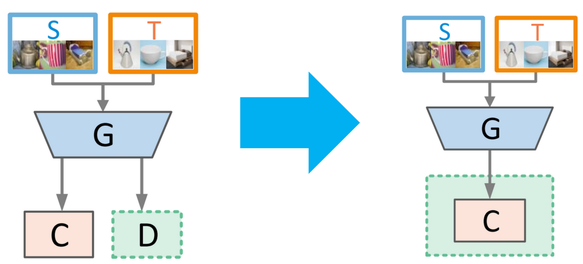


Discriminator-based Unsupervised Domain Adaptation (UDA)

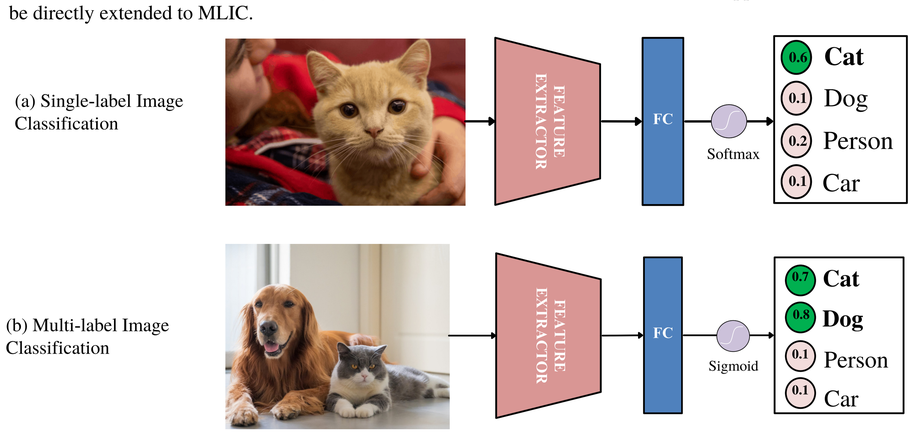


An additional domain discriminator leads to the problem of mode collapse.

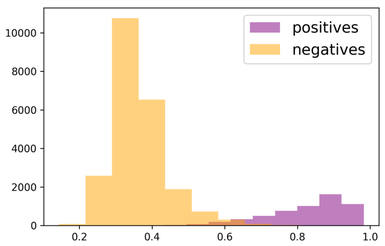
Discriminator-free Unsupervised Domain Adaptation for Single-label Image Classification



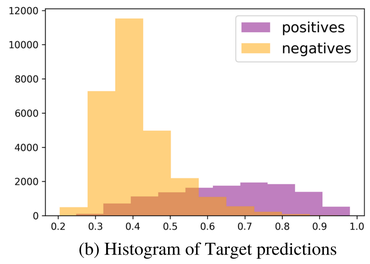
Class correlations for adversarial critic, to reuse the classifier as a discriminator. Similar approach can not be directly extended to MLIC.



Intuition: Histograms of Probability Predictions



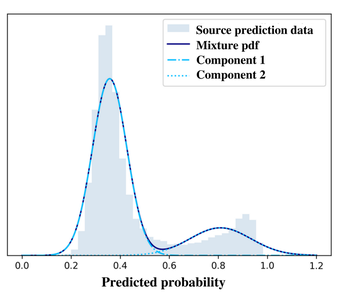
(a) Histogram of Source predictions

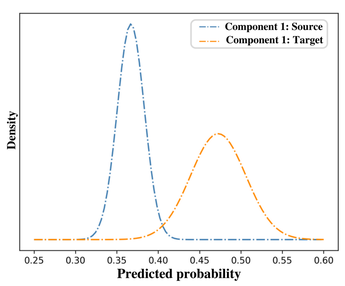


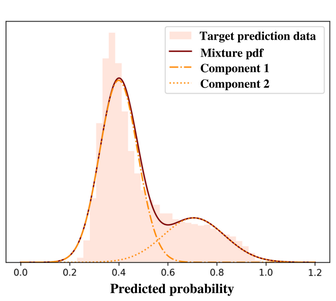
(b) Histogram of Target predictions

The distribution of the probability predictions can be leveraged for defining a suitable critic.

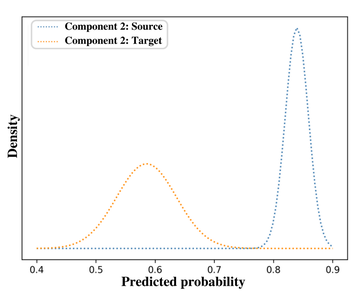
An Implicit Discriminator-free Adversarial Learning Strategy for Multi-lable Image Classification







(a) GMM fitting



(b) Gaussians of components

A two component Gaussian Mixture Model (GMM) can be fitted on Source and Target predictions. A Fréchet distance between the Gaussian pairs of source and target is used as adversarial critic.

$$d_{\mathrm{F}}^2(\mathcal{N}(z_1|\mu_1,\sigma_1),\mathcal{N}(z_2|\mu_2,\sigma_2)) = (\mu_1-\mu_2)^2+(\sigma_1-\sigma_2)^2$$

Experiments

Cross-sensor domain shift

Type	Method	# params (millions)	mean Average Precision (%)			
			AID → UCM	UCM → AID	AID → DFC	UCM → DFC
MLIC	ResNet101	42.5	57.5	51.7	56.9	66.4
	ML-GCN	44.9	53.7	51.3	58.9	64.6
	ML-AGCN	36.6	55.2	52.1	51.6	70.3
	ASL (TResNetM)	**29.4**	55.4	54.1	56.1	68.9
Disc-based	DANN (TResNetM + ASL)	**29.4**	52.5	51.6	43.0	64.1
	DA-MAIC (TResNetM+ASL)	36.6	54.4	50.5	55.4	65.8
Disc-free	DALN (TResNetM + ASL)	**29.4**	53.1	53.2	44.7	65.6
	DDA-MLIC (OURS)	**29.4**	**63.2**	**54.9**	**62.1**	**70.6**

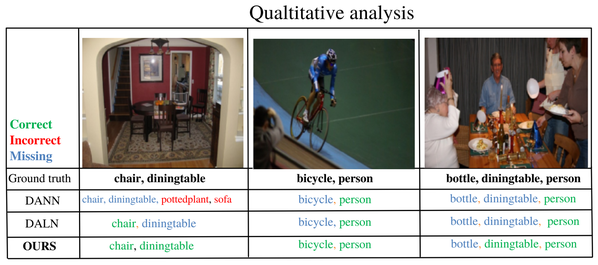
Sim2Real domain shift

Type	Method	# params (millions)	mean Average Precision (%)	
			VOC --> Clipart	Clipart --> VOC
MLIC	ResNet101	42.5	38.0	50.1
	ML-GCN	44.9	43.5	43.1
	ML-AGCN	36.6	53.7	38.0
	ASL (TResNetM)	**29.4**	56.8	64.2
Disc-based	DANN (TResNetM + ASL)	**29.4**	47.0	67.0
	DA-MAIC (TResNetM+ASL)	36.6	**62.3**	74.3
Disc-free	DALN (TResNetM + ASL)	**29.4**	45.0	66.7
	DDA-MLIC (OURS)	**29.4**	61.4	**77.0**

Cross-weather domain shift

Type	Method	# params (millions)	mean Average Precision (%)
			Cityscapes --> Foggy
MLIC	ResNet101	42.5	58.2
	ML-GCN	44.9	56.6
	ML-AGCN	36.6	60.7
	ASL (TResNetM)	**29.4**	61.3
Disc-based	DANN (TResNetM + ASL)	**29.4**	53.5
	DA-MAIC (TResNetM+ASL)	36.6	61.9
Disc-free	DALN (TResNetm + ASL)	**29.4**	54.8
	DDA-MLIC (OURS)	**29.4**	**62.3**

Qualtitative analysis



Acknowledgement
This research was funded in whole, or in part, by the Luxembourg National Research Fund (FNR), grant references BRIDGES2020/IS/14755859/MEET-A/Aouada and BRIDGES2021/IS/16353350/FaKeDeTeR. For the purpose of open access, and in fulfillment of the obligations arising from the grant agreement, the author has applied a Creative Commons Attribution 4.0 International (CC BY 4.0) license to any Author Accepted Manuscript version arising from this submission.

Github

Contact Us