G²NetPL: Generic Game-Theoretic Network for Partial-Label Image Classification

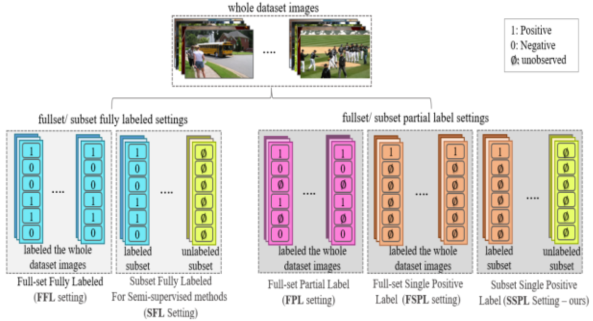
Rabab Abdelfattah[1] Xin Zhang[1] Mostafa M. Fouda[2] Xiaofeng Wang[1] Song Wang[1]

University of South Carolina, USA Idaho State University, USA

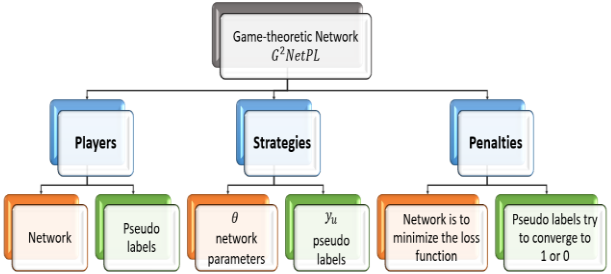
BMVC 2022

Motivation

- Multi-label image classification aims to predict all possible labels in an image which it is **expensive** to annotate all the labels.
- To relieve the annotation burden of full labeling, **partial-label learning** is used.

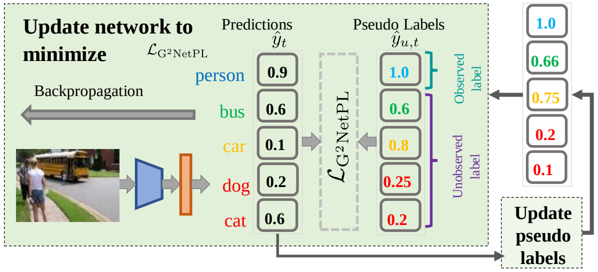


Basic elements of Game



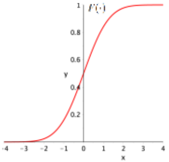
Overview of G^2NetPL

In G^2NetPL, each unobserved label is associated with a soft pseudo label, which, together with the network, formulates a two-player non-zero-sum non-cooperative game.



Loss Functions

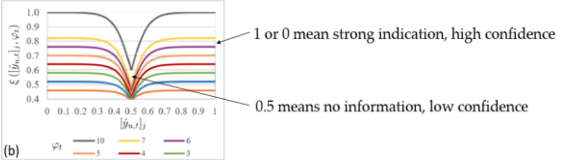
$$\mathcal{L}_{ACE}(\hat{y}_t, y_{u,t}) = \sum_{j=1}^{L}[\mathcal{L}([\hat{y}_t]_j, F([y_{u,t}]_j)) + \lambda_j F([y_{u,t}]_j)(1 - F([y_{u,t}]_j))]$$



$$\mathcal{L}_{G^2NetPL} = \mathcal{L}_{obs} + \mathcal{L}_{unobs}.$$

$$\mathcal{L}_{unobs} = \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{U}_i} \xi([\hat{y}_{u,t}]_j, \varphi_t) \mathcal{L}([\hat{y}_t]_j, [\hat{y}_{u,t}]_j)$$

The pseudo labels will gradually build up their confidence during iterations.

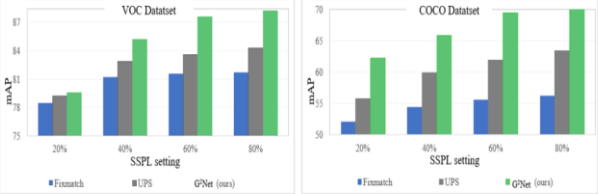


Experiments

Table 1: Quantitative results (mAP) of multi-label image classification on four different datasets. Bold represents the highest mAP and underline represents the second-best among FSPL setting (Single positive and No negative).

Losses	Observed		End-to-End Setting		
	Positive	Negative	VOC	COCO	NUS
$\mathcal{L}_{BCE}$[25]	All	All	89.1	75.5	52.6
$\mathcal{L}_{BCE-LS}$	All	All	90.0	76.8	53.5
$\mathcal{L}_{AN}$ [18]	Single	No	85.1	64.1	42.0
$\mathcal{L}_{AN-LS}$ [7]	Single	No	86.7	<u>66.9</u>	44.9
$\mathcal{L}_{WAN}$ [22]	Single	No	86.5	64.8	<u>46.3</u>
$\mathcal{L}_{EPR}$ [7]	Single	No	85.5	63.3	46.0
$\mathcal{L}_{ROLE}$ [7]	Single	No	<u>87.9</u>	66.3	43.1
$\mathcal{L}_{G^2Net}$ **(ours)**	Single	No	**88.8**	**72.4**	**49.7**

Comparison with Semi-supervised models:



Convergence of pseudo labels during the epochs: