Collaborative Unsupervised Visual Representation Learning From Decentralized Data

Weiming Zhuang[1,2], Xin Gan[1], Yonggang Wen[1], Shuai Zhang[2], Shuai Yi[2]

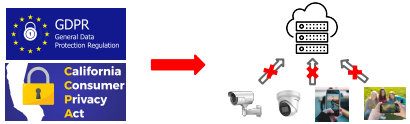
2021 ICCV OCTOBER 11-17 VIRTUAL

Unsupervised Representation Learning

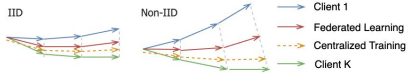
- Learn visual representations without labels
- Achieved remarkable performance on **centralized data** available on the Internet

Challenges

- Unable to centralize growing **decentralized** unlabeled data due to **privacy concerns**



- Non-independently and identically distributed (**non-IID**) data leads to **divergence** [1]



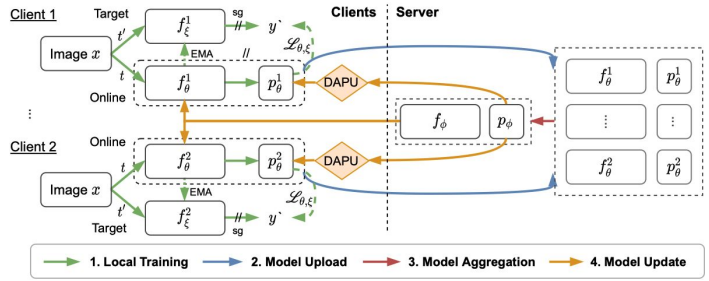
Contributions

- FedU, a new federated unsupervised representation learning framework
- Address non-IID problem:
  - Design a communication protocol
  - Propose divergence-aware predictor update

Existing works

1. Single client training: bad result.
2. Federated learning: need labels
3. Federated unsupervised learning:
   - [2]: bypass non-IID problem
   - [3]: potential privacy leakage risk

Proposed: FedU Framework



Iterate four key activities until stopping conditions

1. Local training: clients conduct unsupervised learning (BYOL [4])
2. Model upload: clients upload trained models to the server
3. Model aggregation: server aggregates them to obtain a new global model
4. Model update: server updates clients' local models with global model

Communication Protocol

Q: Which encoder to aggregate?
A: Online encoder -- has latest learnt knowledge

Q: Which encoder to update?
A: Online encoder -- keep target encoder for regression targets

Divergence-aware Predictor Update

Q: Which predictor to update?

$$p_\theta = \begin{cases} p_\phi & \left\|\theta^r - \phi^{r-1}\right\|_2^2 < \mu \\ p_\theta & \text{otherwise} \end{cases}$$

Intuition:

- Small divergence: use the global p_θ
- Large divergence: use the local p_Φ

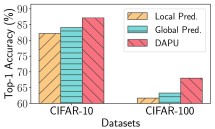
Evaluation

- Setup: 5 clients, one fifth of total classes per client

Method	Architecture	Param.	CIFAR-10		CIFAR-100	
			IID	Non-IID	IID	Non-IID
Single client training	ResNet-18	11M	81.24	71.98	51.33	49.69
Single client training	ResNet-50	23M	83.16	77.84	57.21	55.16
FedSimCLR [1] [36]	ResNet-50	23M	68.10	64.06	39.75	38.70
FedCA [36]	ResNet-50	23M	71.25	68.01	43.30	42.34
FedSimSiam [2]	ResNet-50	23M	79.64	76.70	46.28	48.80
FedU (ours)	ResNet-18	11M	85.21	78.71	56.52	57.08
FedU (ours)	ResNet-50	23M	**86.48**	**83.25**	**59.51**	**61.94**

- Linear evaluation

Aggregate	Update	Accuracy (%)	
		Global Pred.	Local Pred.
Online	**Online**	**84.07**	**82.18**
Online	Target	9.99	19.22
Online	Both	81.24	18.23
Target	Online	82.10	78.06
Target	Target	9.99	25.02
Target	Both	82.32	29.03



- Left: Ablation on the communication protocol
- Right: Ablation on DAPU (Divergence-aware Predictor Update), compared with using only global or local predictors

References

[1] Zhao, Yue, et al. "Federated learning with non-iid data." arXiv preprint arXiv:1806.00582 (2018).

[2] van Berlo, et al. "Towards federated unsupervised representation learning." In Proc. EdgeSys, 2020.

[3] Zhang, Fengda, et al. "Federated unsupervised representation learning." arXiv preprint arXiv:2010.08982 (2020).

[4] Grill, Jean-Bastien, et al. "Bootstrap your own latent: A new approach to self-supervised learning." In Proc. NIPS, 2020

Paper & Code: https://weiming.me/