CLOT: Contrastive Learning-Driven and Optimal Transport-Based Training for Simultaneous Clustering

Mohammed Aburidi & Roummel Marcia
Department of Applied Mathematics, University of California Merced, CA, USA

Introduction

Problem description: Clustering via representation learning is one of the most promising approaches for self-supervised learning of deep neural networks. It aims at obtaining artificial supervisory signals from unlabeled data.

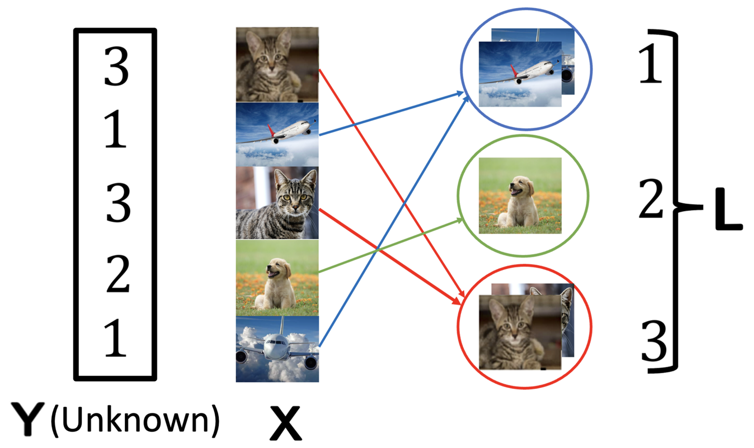


Fig. 1: Labels assignment using optimal transport.

Motivation: Learning a powerful representation often requires a large-scale dataset with manually curated ground-truth labels, which has proven to be a bottleneck for the continued development of state of the arts performance and in its deployment in many application areas.

Method: We propose a deep-based clustering method called Contrastive Learning driven and Optimal Transport-based (CLOT) clustering which focuses on the problem of obtaining the labels simultaneously.

Results: We test our framework on three standard benchmarks: CIFAR-100, ImageNet-10 and STL-10. Our framework outperforms eight state-of-the-art methods on all three datasets.

Materials & Methods

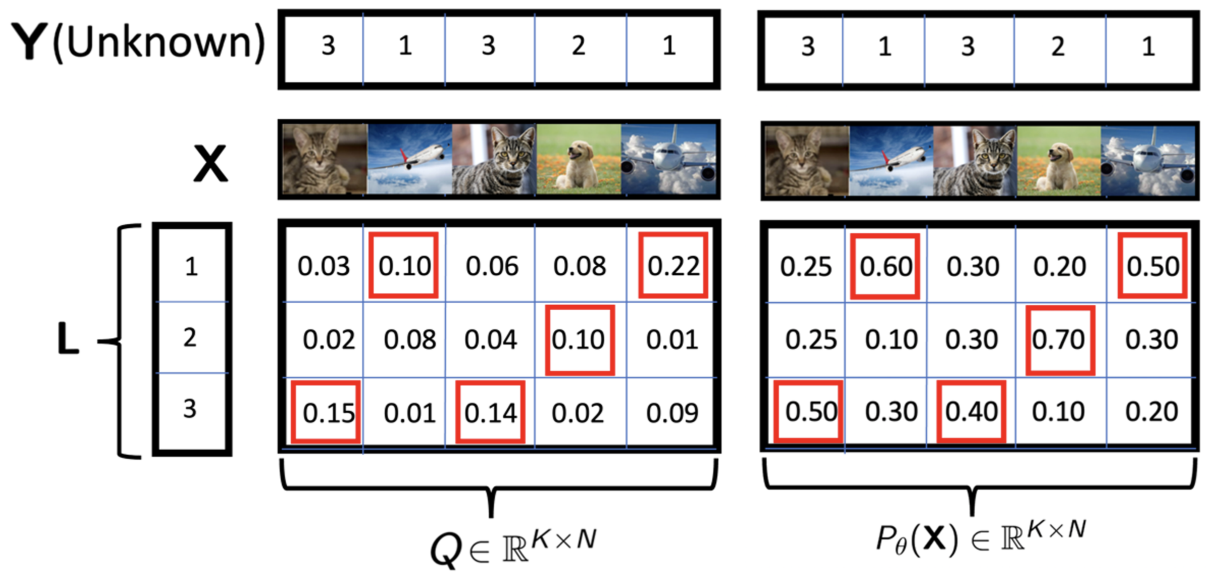
Let:

- $\mathbf{X}$ be a batch of unlabeled images (Source).
- $\mathbf{L}$ be the set of K cluster labels (Target).
- $\mathbf{Y} = \{y_i\}_{i=1}^N$ be the unknown true labels.
- $Q \in \mathbb{R}^{K\times N}$ be the label assignment matrix.
- $P_\theta(\mathbf{X}) \in \mathbb{R}^{K\times N}$ be the predicted probability matrix

Cross-entropy objective:

$$\mathcal{L}_{opt}(P_\theta, Q) = -\sum_{i=1}^{N}\sum_{j=1}^{K} Q(y=j|x_i) \log P_\theta(y=j|x_i)$$

where: $[Q(\mathbf{X})]_{ij} = Q(y=j|x_i)$ $\quad [P_\theta(\mathbf{X})]_{ij} = P_\theta(y=j|x_i)$



Alternating minimization proposed approach:
Step 0: Randomly initialize θ, and compute P_θ.
Step 1: Given P_θ, find label assignments Q by solving the OT.
Step 2: Given Q, optimize the model parameters θ and compute P_θ.

We add equality constraints to avoid degeneracy (assigning all images to one class)
Using matrix notation, the previous optimization problem can be written as

$$\underset{Q}{\text{minimize}} \quad \langle Q, -\log P\rangle_F$$

$$\text{subject to} \quad Q\mathbf{1}_N = \frac{1}{K}\mathbf{1}_K \qquad Q^T\mathbf{1}_K = \frac{1}{N}\mathbf{1}_N$$

Applying the entropic regularization, our problem is given by

$$\underset{Q\in\mathbf{Q}}{\text{minimize}} \quad \langle Q, -\log P\rangle - \frac{1}{\lambda}S(Q)$$

Using Sinkhorn-knopp algorithm, the minimizer is

$$Q^* = \text{Diag}(u)P^\lambda \text{Diag}(v)$$

Model improvements:

A. Image transformation-invariant model

B. Combine contrastive learning: additional multi-layer perceptron (MLP) to obtain feature vectors $\mathbf{z}^a$ and $\mathbf{z}^b$.

Results 1: Classification

Datasets: CIFAR-100, STL-10, and ImageNet-10. Each dataset contains 10 classes except CIFAR-100, which contains 20 classes.

Evaluation Metrics: Accuracy (ACC), Normalized Mutual Information (NMI), and Adjusted Rand Index (ARI).

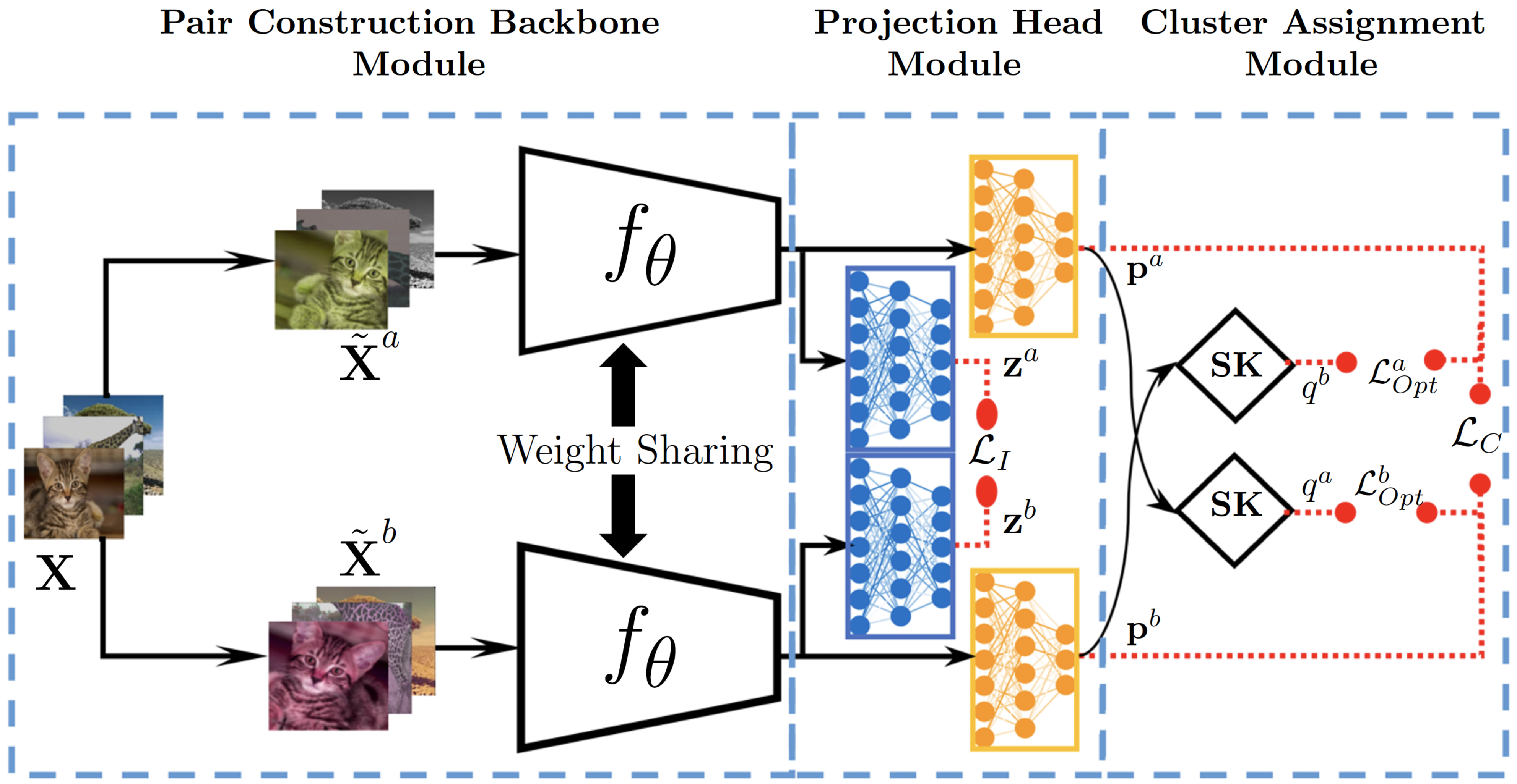


Fig. 3: CLOT clustering framework

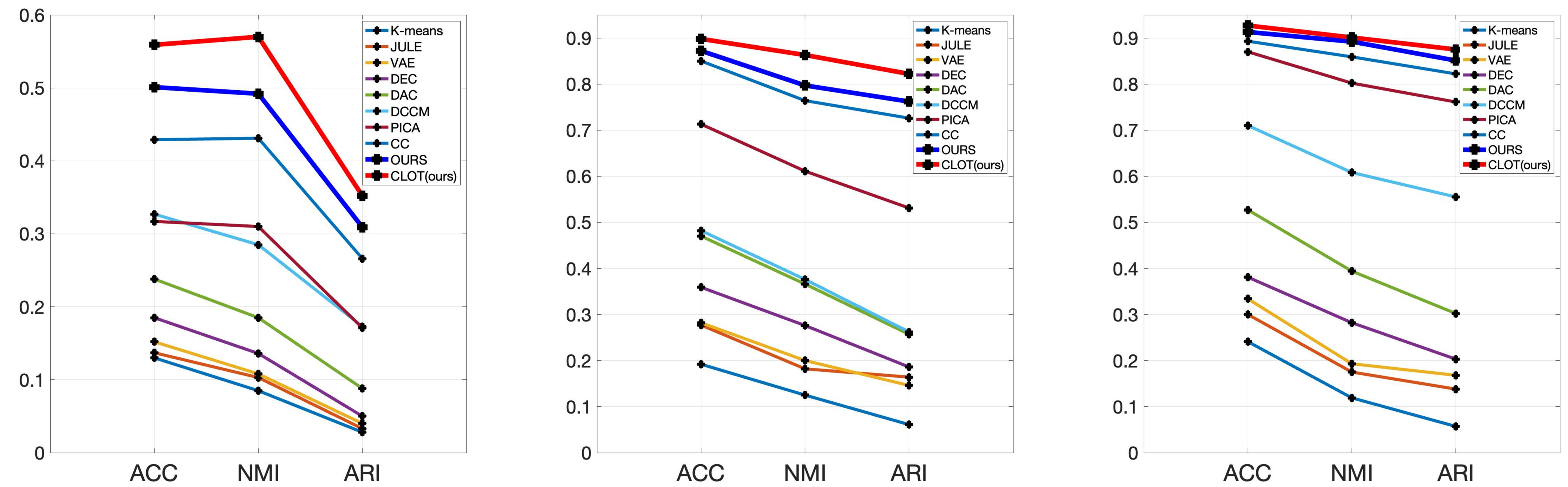


Fig. 4: The clustering performance on three image benchmarks, CIFAR-100, STL-10 and ImageNet-10 from left to right.

Conclusion

- We present an online clustering method that is based on counteractive feature representation learning and contrasting cluster assignments
- Compared to existing state-of-the-art methods, the proposed CLOT shows promising performance in clustering on three challenging datasets

References

[1] Gabriel Peyré and Marco Cuturi. Computational optimal transport, 2020.

[2] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations, 2020.

Contact Information

Email rmarcia@ucmerced.edu