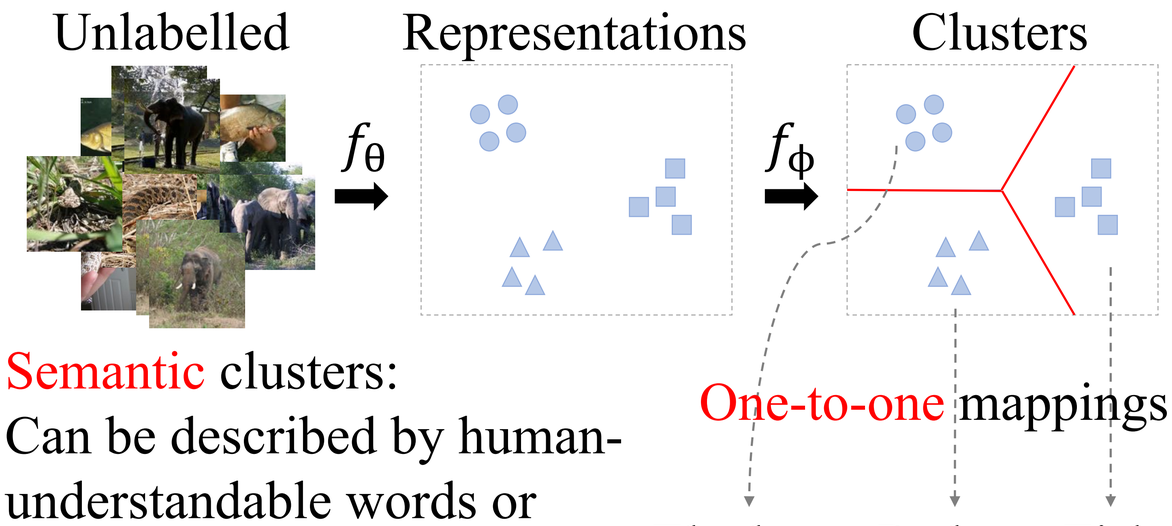
Deep Clustering by Semantic Contrastive Learning

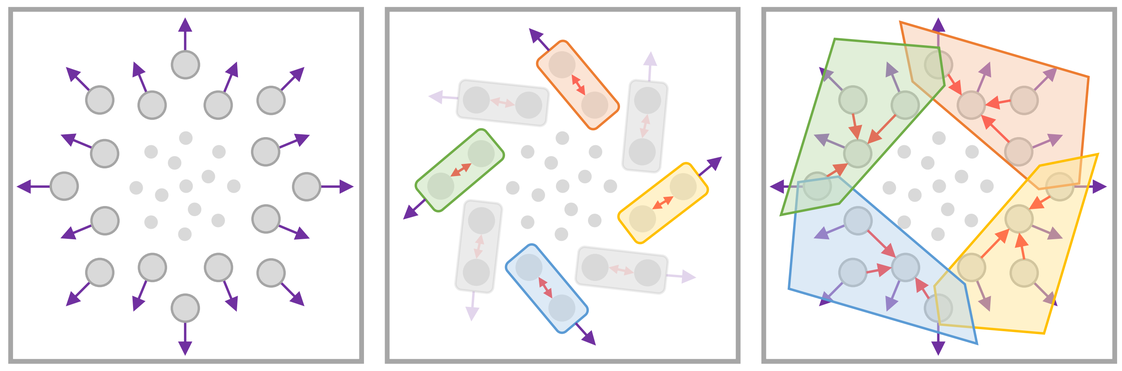
Jiabo Huang jiabo.huang@qmul.ac.uk
Shaogang Gong s.gong@qmul.ac.uk
Queen Mary University of London, London, E1 4NS, UK

1 Deep clustering of Unlabelled Images



Semantic clusters:
Can be described by human-understandable words or phrases, e.g. class labels

2 Related works & Limitations

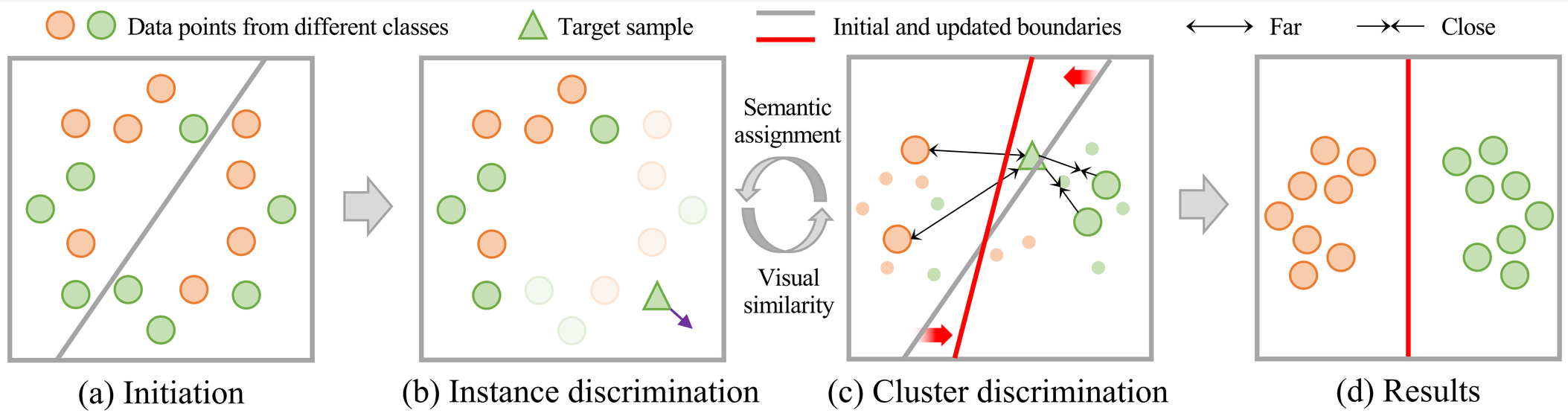


(a) Sample-specific (b) Neighbourhood (c) Joint Clustering

(a) Insensitive to potential class memberships
(b) Restricted to subtle intra-class variations
(c) Instance contrast v.s. clustering - contradictory

3 Overview

(a) Start from randomly initial representations and decision boundaries



(b) Pulling each instance away from only its pseudo-negative samples of other clusters

(c) Simultaneously learning decision boundaries according to the distance-based cluster structure

4 Cross-cluster instance discrimination

➢ **Semantic Memory** $\mathcal{M} = \{M_1, M_2, \dots, M_C\}$

Maintaining C independent memory banks with each corresponding to a target cluster

➢ **Per-instance contrastive set** Q_i

For sample $\boldsymbol{I}_i$ with a pseudo label y_i, its contrastive set Q_i is composed of its pseudo negative samples

$$Q_i = \{\widetilde{\boldsymbol{k}} | \widetilde{\boldsymbol{k}} \in M_j \; \forall \, j \in [1, C] \text{ and } j \neq y_i\}$$

➢ **Instance discrimination objective**

To identify samples' feature $\boldsymbol{q}_i$ and their perturbed copies $\boldsymbol{k}_i$ from their contrastive set

$$\mathcal{L}_{\mathrm{ID}}(\boldsymbol{I}_i) = -\log\left(\frac{\exp(\cos(\boldsymbol{q}_i, \boldsymbol{k}_i)/\tau)}{\sum_{\widetilde{\boldsymbol{k}} \in Q_i \cup \{\boldsymbol{k}_i\}} \exp(\cos(\boldsymbol{q}_i, \widetilde{\boldsymbol{k}})/\tau)}\right)$$

5 Online cluster discrimination

➢ **Distance-based cluster structure**

Taking samples in each semantic memory bank M_i as the anchors of the corresponding cluster

$$\tilde{p}_{i,j} = \frac{\sum_{\widetilde{\boldsymbol{k}} \in M_j} \exp(\cos(\boldsymbol{q}_i, \widetilde{\boldsymbol{k}})/\tau)}{\sum_{j'=1}^{C} \sum_{\widetilde{\boldsymbol{k}} \in M_{j'}} \exp(\cos(\boldsymbol{q}_i, \widetilde{\boldsymbol{k}})/\tau)}$$

➢ **Decision boundaries**

Modelling decision boundaries by an FC layer

$$\boldsymbol{p}_i = \mathrm{Softmax}(\mathrm{W}^\top \boldsymbol{q}_i + B) \in \mathcal{R}^C$$

➢ **Cluster discrimination objective**

Updating boundaries to yield consistent $\boldsymbol{p}_i$ with $\tilde{\boldsymbol{p}}_i$

$$\mathcal{L}_{\mathrm{CD}}(\boldsymbol{I}_i) = \sum_{j=1}^{C} -\tilde{p}_{i,j} \log p_{i,j}$$

6 Experiments

➢ **Deep clustering**

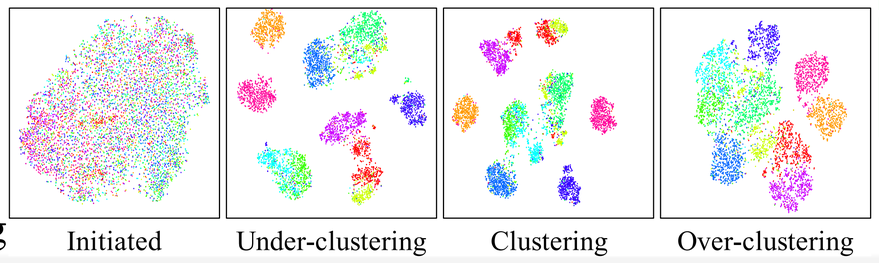
	CIFAR100	IN-Dogs*	Tiny-IN*
DCCM	0.327	0.383	0.108
PICA	0.337	0.352	0.098
GAT	0.281	0.322	-
CC†	0.429	0.429	0.140
GRLC†	0.425	0.484	-
GCC†	0.472	0.526	0.138
SCL†	**0.482**	**0.763**	**0.172**

IN*: ImageNet; Method†: w/ contrastive learning

➢ **Representation Learning**

	CIFAR10	CIFAR100	STL10
MoCo	0.528	0.360	0.561
PAD	0.626	0.288	0.465
DeepCluster	0.374	0.189	0.334
SCL	**0.813**	**0.482**	**0.638**

➢ **Feature visualisation**



➢ **Visual case examples**