Week 16

▷ Exploring Balanced Feature Spaces For Representation Learning

Supervised learning $\xrightarrow{\times}$ Imbalanced dataset

Self-supervised learning $\xrightarrow{?}$ Imbalanced dataset

↓

learn data representations without requiring manual annotations

leverage the vast amount of unlabeled data in the wild to obtain strong representation models

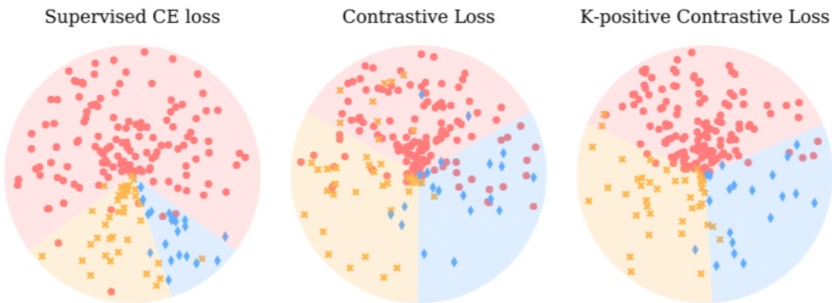

Figure 1: *Feature spaces learned with different losses given an imbalanced dataset.* The supervised cross-entropy (CE) learns a space biased to the dominant class. The space learned by unsupervised contrastive loss is balanced but less semantically discriminative. Our proposed k-positive contrastive loss learns a balanced and discriminative feature space. The shadow area ( ) indicates the decision boundary of each class.

Using the contrastive loss can obtain representation models generating a balanced feature space.

— Balanced Feature Spaces From Contrastive Learning

Representation models (imbalanced training datasets) $\begin{cases} \text{Self-supervised} \\ \text{supervised} \end{cases}$

Representation model $f_\theta : \underset{\text{Sample}}{x_i} \to v_i \in \underset{\text{feature space}}{V}$

Dataset for training $D_{rep\text{-}train} = \{x_i, y_i\}$, $i = 1, 2, \dots, N$

Discrete instance distribution $\{q_1, q_2, \cdots, q_C\}$ with $q_j = n_j / N$, C classes

Final classification prediction $\hat{y} = \arg\max[W^T v + b]$

Representation learning methods

(1) supervised cross-entropy (CE) loss:

$$\mathcal{L}_{CE} = \frac{1}{N}\sum_{i=1}^{N} -\log p_{y_i}, \quad \text{where } p_{y_i} = \text{softmax}(W_{y_i}^T v_i + b)$$

Have strong semantic discrimination ability but its generated feature space is easily biased by the imbalance of the training instance distribution.

(2) Contrastive loss (CL) → semantic free:

$$\mathcal{L}_{CL} = \frac{1}{N}\sum_{i=1}^{N} -\log \frac{\exp(v_i \cdot v_i^+ / \tau)}{\exp(v_i \cdot v_i^+/\tau) + \sum_{v_i^- \in V^-} \exp(v_i \cdot v_i^- / \tau)}$$

Balancedness of feature spaces

V is balanced if the $\{v_i\}$ from different classes within it have similar degrees of linear separability.

↓

$a_i, i = 1, 2, \cdots C$, the accuracy of a linear classifier

$$\beta(V) \triangleq \frac{1}{C^2}\sum_{i,j}^{C} \exp\left(-\frac{|a_i - a_j|^2}{\sigma}\right), \quad \text{where } a_j = \frac{\#\{v_i \mid \hat{y}_i = j, y_i = j, v_i \in V\}}{\#\{v_i \mid y_i = j, v_i \in V\}}$$

if all the class-wise accuracies are equal, maximize $\beta(V)$

Experimental protocol

Imbalanced

(1) Representation learning: pre-train f_θ on $D_{\text{rep-train}}$ using $\mathcal{L}_{CE}$ and $\mathcal{L}_{CL}$

(2) Classifier learning: train (W, b) on top of f_θ on D_{train}

(3) Representation evaluation: $D_{\text{test}} \xrightarrow{f_\theta} \beta(V)$

} balanced

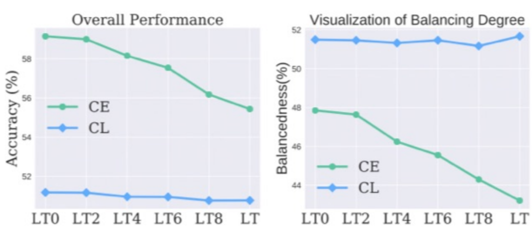

Figure 2: Classification accuracy (left) and balancedness (right) of the representations learned from cross-entropy (CE) loss and contrastive loss (CL) on datasets (LT0 to LT) with increasing imbalance.

The model trained with $\mathcal{L}_{CL}$ generates a more balanced feature space.

K-Positive Constrastive Loss

Embed semantic discriminativeness into the representations while maintaining desired balancedness.

$$\mathcal{L}_{KCL} = \frac{1}{N(k+1)} \sum_{i=1}^{N} \sum_{v_j^+ \in \{\tilde{v}_i\} \cup \underline{V_{i,k}^+}} -\log \frac{\exp(v_i \cdot v_j^+ / \tau)}{\exp(v_i \cdot \tilde{v}_i / \tau) + \sum_{v_j \in V_i} \exp(v_i \cdot v_j / \tau)}$$

k instances from the same clas

learn representations with stronger discriminative ability

use the same number of instances (k) for all the classes (balance the learned feature space)

▷ Questions

1. In our task, we also need to map the image into the representation space. It's generally realized by unsupervised contrastive learning. And according to the above paper, contrastive learning can learn a balanced feature space from imbalanced datasets.

   Does it mean that this problem is irrelevant to the balanceness assumption of the original dataset ?

2. KCL combines semantic label of data (supervised), can we apply similar idea to improve the balanceness and separability of feature space ?

3. Our task aims to select instances for labeling in a balanced way, we don't know how many classes there are, we only know the annotation budget.

   Is it an unsupervised classification problem?

   What if the annotation budget cannot guarantee the balanceness of each class ?