Open world semi supervised learning

Setting:

$\in C_\ell$

$$D_\ell = \{(x_i, y_i)\}_{i=1}^{N}$$

C_u

$$D_u = \{(x_i)\}_{i=1}^{m}$$

Dataset Assumption: Category / class set:

$$\boxed{C_\ell \cap C_u \neq \emptyset}$$ Different from NCD

$$C_\ell \neq C_u$$

$$C_s = C_\ell \cap C_u$$ // Seen classes.

$$C_u = C_u \setminus C_\ell$$ // unseen/Novel classes.

Notations:

$$f_\theta : \mathbb{R}^N \to \mathbb{R}^D$$

labeled

feature: $$Z_\ell = \{z_i \in \mathbb{R}^D\}_{i=1}^{n}$$

$$Z_u = \{z_i \in \mathbb{R}^D\}_{i=1}^{m}$$

classification head weight matrix.

$$W : \mathbb{R}^D \rightarrow \mathbb{R}^{|C_l \cup C_u|} \rightarrow \text{a softmax.}$$

final prediction: $c_i = \underbrace{\arg\max (w^T \cdot z_i)}_{\text{one of the } C_l \cup C_u \text{ classes.}}$

for unknown no of classes we initialize the head with Large Number of classes. [OVER CLUSTER solution]

Final optimization Objective:

pairwise objective

$$\boxed{\mathcal{L} = \mathcal{L}_s + \eta_1 \mathcal{L}_p + \eta_2 \mathcal{R}}$$

supervised

Regularizer.

① Supervised Objective with Uncertainty Adaptive margin $\mathcal{L}_s$

- utilization of categorical annotated labels.

ith row.

usual case

$$\mathcal{L}_s = \frac{1}{n} \sum_{z_i \in \mathcal{Z}_l} -\log \left[\frac{e^{w_{y_i}^T \cdot z_i}}{e^{w_{y_i}^T \cdot z_i} + \sum_{j \neq i} e^{w_{y_j}^T \cdot z}} \right]$$

- result in learning a classifier with larger margin.
- Result model bias towards the known classes [smaller intra-class variance for the known compared to novel classes]

Introduce uncertainty Adaptive margin

[propose a normalization!]

$$\boxed{\mathcal{L}_s = \frac{1}{n} \sum_{z_i \in \mathcal{Z}_l} -\log \left[\frac{e^{s(w_{y_i}^T \cdot z_i + \lambda \bar{u})}}{e^{s(w_{y_i}^T \cdot z_i + \lambda \bar{u})} + \sum_{j \neq i} e^{s\, w_{y_j}^T \cdot z_i}} \right]}$$

uncertainty $\bar{u}$ → how?

scale s

$$\bar{u} = \frac{1}{|D_u|} \sum_{x \in D_u} var(y \mid x = x)$$

$$= \frac{1}{|D_u|} \sum_{x \in D_u} Pr(y = 1 \mid x)\, Pr(y = 0 \mid x)$$

$$\approx \underbrace{\frac{1}{|D_u|} \sum_{x_i \in D_u} 1 - \max Pr(y = k \mid x = x_i)}_{\text{Group uncertainty.}}$$

Requires
$$\begin{cases} z_i = \dfrac{z_i}{|z_i|} \\[2ex] w_j = \dfrac{w_j}{|w_j|} \end{cases}$$

(ii) Pairwise Objective

$$\mathcal{L}_p = \frac{1}{m+n} \sum_{z_i,\, z_j \in \{z_l \cup z_u,\ \underbrace{z' \cup z'_u}_{\text{Augmentation}}\}} - \log \overbrace{\langle \underset{\downarrow \text{softmax}}{\sigma}(w^T \cdot z),\ \sigma(w^T \cdot z') \rangle}^{\text{Dot product}}$$

works with positive only !!

(iii) Regularizer term

$$\mathcal{R} = KL\left(\frac{1}{m+n}\sum_{z_i \in Z_r \cup Z_u} \sigma(w^T \cdot z_i) \,\|\, \mathcal{P}(y)\right)$$