Divide and conquer:NCD

Notations:

Baseset (labeled) : $\mathcal{D}^b = \{(x_i, y_i)\}_{i=1}^{N^b}$

Novel set (unlabeled) : $\mathcal{D}^n = \{x_j\}_{j=N^b+1}^{N^n+N^b}$

$\hookrightarrow$ number of classes C^N (known)

Encoder, $z = \phi(x)$

Objective:

mapping Image Space, $\mathcal{X} = \{x_i\}_{i=1}^{N_b+N_c} \rightarrow \mathcal{Y} = \{l\}_{l=1}^{C_b+C_n}$

Compositional Experts:

Batchwise Experts: A linear classifier $\phi^b(\cdot)$

Prediction output: $\boxed{\hat{y}_i^{\phi^b} = \phi^b(z_i)}$; $z_i \in \mathcal{D}^b$

$\hookrightarrow$ Normalized

Novel batch expert: Another projection layer $\phi^n(\cdot)$

$\rightarrow z_j^{\phi}$ (Normalized)

$$\boxed{\hat{y}_i^{\phi^n} = \phi^n(\phi^r(x_i))} \quad ; x_i \in D^n$$

Cross entropy loss:

$$\mathcal{L}_{CE}(\hat{y}^\phi, y) = -y \log \sigma(\hat{y}^\phi / c)$$

[utilize the labeled & apply pseudo label for the rest]

Regularization to avoid class leakage:

$$\mathcal{L}_{reg}(\hat{y}^\phi) = \left(\sum_{c \in \hat{y}} (\hat{y}_c^\phi)^2 \right)^{1/2}$$

$, — C^n + C^b$

(kinda entropy loss)

Encourage minimum entropy

Classwise Expert:

$\hat{y}^{\phi^b} = \phi^b(x_i)$ [only base class] concat

$$y^{\psi^n} = \psi^n(\psi'^n(z_i))$$ [only expert class]

$$\boxed{\mathcal{L}_{CE}\left(\left[\hat{y}^{\psi^b}, \hat{y}^{\psi^n}\right], y\right)}$$

Training target:

pseudo labeling: Equally partition the novel clusters. ψ^n, ϕ^n

Batch of novel samples:

$$\boxed{Z = \left[z_1^{\phi}, z_2^{\phi}, \dots z_{B^n}^{\phi}\right] \in \mathbb{R}^{B^n \times d}}$$

cluster center: $W = \left[w_1, w_2 \dots w_{C^n}\right] \in \mathbb{R}^{C^n \times d}$

labels

embedding

Goal $\Rightarrow \max_{Y \in P} \operatorname{tr}\left(Y W Z^T\right) + \epsilon H(Y)$

smoothness.

$$\left[y_1^{\phi^n}, y_2^{\phi^n}, \ldots y_{B^n}^{\phi^n} \right] \in \mathbb{R}^{B^n \times C^n}$$

Solution by Sinkhoren-knopp algorithm.

$\boxed{\text{mean of } z^{\phi}, z^{\psi}}$

Local aggregation:

$$y^n = \alpha \frac{y^{\phi^n} + y^{\psi^n}}{2} + (1-\alpha) \sum_{k=1}^{N^q} s_k y_k^q$$

$$s_k = \frac{\exp\left(\bar{z} \cdot z_k^q / \tau\right)}{\sum_{k=1}^{N^q} \exp\left(\bar{z} \cdot z_k^q / \tau\right)}$$

$\hookrightarrow$ Attention mechanism

[N^q novel sample in the dictionary]