Ncd with ranking loss

unlabeled dataset:

$$D^u = \{ x_i^u, i = 1 \dots m \}; x_i^u \in \mathbb{R}^{3 \times H \times W}$$

cluster into classes C^u (known)

labeled dataset:

$$D^l = \{ (x_i^l, y_i^l) ; i = 1 \dots N \}$$

$$y_i^l \in \{ 1, \dots C^l \}$$

image representation: $\phi : x \mapsto \phi(x) \in \mathbb{R}^d$??

3 ideas to learn:

Self-supervised: & supervised learning

RotNet

[Rotation prediction]

$$z_i^l = \phi(x_i^l)$$

$$\eta^l : \mathbb{R}^d \to \mathbb{R}^{C^l}$$

$$L_{CE} = -\frac{1}{N} \sum_{i=1}^{N} \log \eta_{y_i}^l (z_i^l)$$

Transfer learning via Rank statistics.

for a unlabeled image pair (x_i^u, x_j^u)

$$z_i^u = \phi(x_i^u)$$

$$z_j^u = \phi(x_j^u)$$

$s_{ij} \in \{0, 1\}$

but how to select it?

Ranking

$$s_{ij} = \mathbb{1}\{\text{top}_k(\phi(x_i^u)) = \text{top}_k(\phi(x_j^u))\}$$

subset of $(1-d)$ top-k element.

New head: $\eta^u : \mathbb{R}^d \to \mathbb{R}^{C^u}$ — dot product.

$$L_{BCE} = -\frac{1}{m^2} \sum_{j=1}^{m} \sum_{i=1}^{m} \Big[s_{ij} \log \eta^u(z_i^u)\, \eta^u(z_j^u) + (1 - s_{ij}) \log\{1 - \eta^u(z_i^u)\, \eta^u(z_j^u)\} \Big]$$

pairwise loss.

Enforcing Prediction to be consistent

$$L_{MSE} = \frac{1}{2}\sum_{i=1}^{n}\left(y_i^1 - \hat{y}_i^1\right) + \frac{1}{m}\sum_{i=1}^{m}\left(y_i^2 - \hat{y}_i^2\right)$$

transformed image.

final loss

$$L = L_{CE} + L_{BCE} + w(t)\, L_{MSE}$$