P PAWS PAWS

unlabeled dataset $\mathcal{D} = (x_i)_{i \in [1, N]}$

Support set $S = \{(x_i, y_i)_{i \in [1, M]}\}$ $M << N$

Leverage both $\boxed{D \text{ \& } S}$

$\boxed{\text{Pretraining } \mathcal{D}, S}$
$\boxed{\text{fine tune with } S}$
twice??

$\boxed{\text{SWAV, BOYL}}$ → Positive only

$x_i \in \mathcal{D}$ ↗ Anchor $\hat{x}_i$ ↘ positive $\hat{x}_i^+$ } minimize cross entropy between them.

Detailed: on hand

$x \in \mathbb{R}^{n \times (3 \times H \times W)}$ → view Anchor

$x^+ \in \mathbb{R}^{n \times (3 \times H \times W)}$ → positive

$x_S \in \mathbb{R}^{m \times (3 \times H \times W)}$
$y_S \in \mathbb{R}^{n \times K}$ } labeled [K total]

→ one hot

Encoder: $\mathbb{R}^{3 \times H \times W} \xrightarrow{f_\theta} \mathbb{R}^d$

→ $z \in \mathbb{R}^{n \times d}$

$z^+ \in \mathbb{R}^{n \times d}$

$z_S \in \mathbb{R}^{S \times d}$] label matrix. → j-th row

similarity Classifier $\pi_d(z_i, z_S) = \sum_{(z_j, y_j) \in z_S} \left(\frac{d(z_i, z_{S_j})}{\sum_{z_{S_k} \in z_S} d(z_i, z_{S_k})} \right) y_j$

similarity matrics: $d(a,b) = \exp\left(\frac{a^T b}{\tau \|a\| \|b\|}\right)$

$$p_i : \pi(z_i; \mathbf{z}_s) = \sigma_\tau(z_i^T \mathbf{Z}_s) \underline{y_s}$$ // softmax prob.

sharpening function $[\rho(p_i)]_k := \dfrac{[p_i]_k^{1/T}}{\sum_{j=1}^{K} [p_i]_j^{1/T}}$; $k = 1 \ldots K$ → temperature.

→ weights sharpening $(p_i)_j$

where, $p_i \in [0,1]^k$

overall objective. for encoder. to minimize

$$\frac{1}{2n} \sum_{i=1}^{n} \left[\underbrace{H(\rho(p^+), p_i)}_{\text{cross entropy}} + \underbrace{H(\rho(p_i), p^+)}_{\text{the heck? why?}} \right] - \underbrace{H(\bar{p})}_{\text{Entropy}}$$

Theoretical Bound:

Assumption: [Balanced class] & [target sharpening & not uniform].

uniform

Prop: Non-collapsing Representation. if rep collapse $\underline{z_i = z \ \forall z_i \in S}$ $d(z, z_i) = d(z, z_j)$.

then $\|\nabla_\theta H(p^+, p)\| > 0$; gradient is positive ?)

Proof: if $d(z_i, z) = d(z_j, z)$

$\Rightarrow p : \pi(z, S) = \frac{1}{n} \sum_{z_i} y_i = \frac{1}{k}$ (yes) // uniform

k ??

p^+ not uniform then, $\|\nabla H(p^+, p)\| > 0$

not uniform (uniform)