Swapping Assignment between multiple view of same images.

SWAV

image x_n

Augmented view x_{nt}

2 standard resolution + V additional low resolution crop.

$\Rightarrow (V+2)$ images.

$z_{nt} = f(x_{nt})$ // after normalization.

code q_{nt} from z_{nt} to a set of

K trainable normal vectors $\underbrace{\{c_1, \ldots c_K\}}_{\text{prototypes.}}$

$\rightarrow$ weight of dense layer!!

z_t, z_s augmented representation

$q_t, q_s \rightarrow$ their corresponding code.

$$L(z_t, z_s) = \underbrace{l(z_s, q_t) + l(z_t, q_s)}_{\text{swapping}}$$

→ overall prototypes.

$$l(z_s, q_t) = -\sum_k q_s^{(k)} \log p_t^{(k)}$$

$$p_t^{(k)} = \frac{\exp\left(\frac{1}{\tau} z_t^T c_k\right)}{\sum_{k'} \exp\left(\frac{1}{\tau} z_t^T c_{k'}\right)}$$

Total loss function:

$$-\frac{1}{N} \sum_{n=1}^{N} \sum_{t \sim T} \left[\frac{1}{\tau} z_{nt}^T C q_{ns} + \frac{1}{\tau} z_{ns}^T C q_{nt} - \log \sum_{k=1}^{K} \exp\left(\frac{z_{nt}^T c_k}{\tau}\right) - \log \sum_{k=1}^{K} \exp\left(\frac{z_{ns}^T c_k}{\tau}\right) \right]$$

$$\begin{cases} C \equiv D \times K \\ z \equiv D \times 1 \\ q \equiv K \times 1 \end{cases}$$

prototypes $c \in D \times 1$

Computing Codes online :

Cluster Assignment :

Assign B samples $[z_1 \ldots z_B]$

to k (prototype) cluster $[c_1, \ldots c_k]$

Requires mapping Q

Equipartition constraint

(to avoid collapse)

Spliting Data uniformly $r = \frac{1}{k} 1_k$

$c = \frac{1}{B} 1_B$

$$Q = \left\{ Q \in \mathbb{R}_+^{k \times B} \;\middle|\; Q 1_B = \frac{1}{k} 1_k ,\ Q^T 1_k = \frac{1}{B} 1_B \right\}$$

[each cluster is selected B/k times]

Cost matrix $:= C^T Z$

[solve this under equipartition constraint]

$M \to D \times D$ cost matrix.

$r \to c$ mapping (PDF)

by transport matrix P

total cost $= \langle P, M \rangle$ // Frobenius product.

$$P \in \{ A \in \mathbb{R}_+^{D \times D} \mid A \mathbb{1}_D = r;\ A^T \mathbb{1}_D = c \}$$

then, $d_M(r, c) = \min_P \langle P, M \rangle$

Optimal transport problem:

regularizer.

$$\min_{Q \in \mathcal{Q}} \langle Q, -C^T Z \rangle + \epsilon H(Q)$$

Solution: $Q^* = \mathrm{Diag}(u) \exp\left(\frac{C^T Z}{\epsilon}\right) \mathrm{Diag}(v)$

$$\left.\begin{array}{l} u \in \mathbb{R}^K \\ v \in \mathbb{R}^B \end{array}\right\} \text{ non negative}$$

computed by small mat mul.

(Sinkhorn-Knopp algorithm)