Episode Based Training (2)

We are going to treat episodes as mini-batches in training

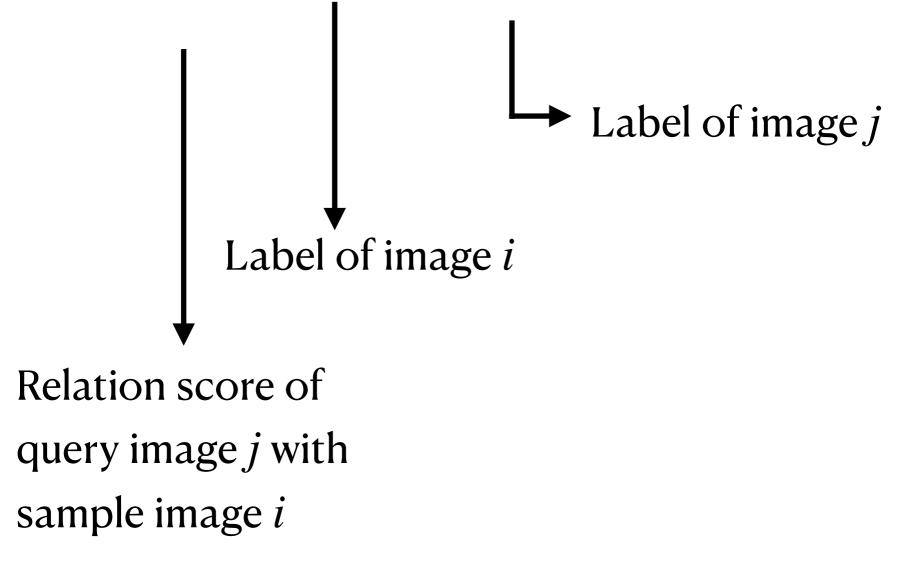
 $\mathcal{L}_{episode}(S, Q) = \sum \sum (r_{i,j} - 1(y_i = y_j))^2$

: loss for an episode defined by the support set S and the query set Q

$$\varphi, \phi \leftarrow \arg\min_{\varphi, \phi} \mathbb{E}_{L \sim \mathcal{T}} \left[\mathbb{E}_{S \sim L, Q \sim L} \left[\mathcal{L}_{episode}(S, Q) \right] \right]$$

sampling a label set over a distribution over possible label sets $\mathcal T$

 $S \sim L$ and $Q \sim L$ represent sampling the support set S and query set Q from L



Episode Based Training (2)

We are going to treat episodes as mini-batches in training

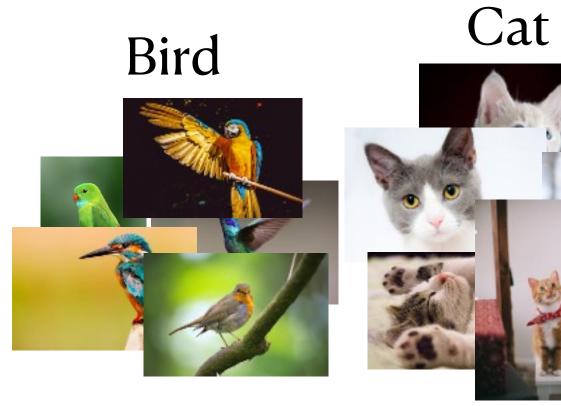
$$\mathcal{L}_{episode}(S,Q) = \sum_{i=1}^{m} \sum_{j=1}^{n} \left(r_{i,j} - 1(y_i = y_j) \right)^2 \text{ :loss for an episode defined by the support set S and the query set Q}$$

 $S \sim L$ and $Q \sim L$ represent sampling the support set S and query set Q from L

$$\varphi, \phi \leftarrow \arg\min_{\varphi, \phi} \mathbb{E}_{L \sim \mathcal{T}} \left[\mathbb{E}_{S \sim L, Q \sim L} \left[\mathcal{L}_{episode}(S, Q) \right] \right]$$

sampling a label set over a distribution over possible label sets $\mathcal T$

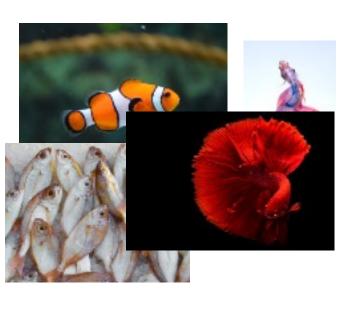




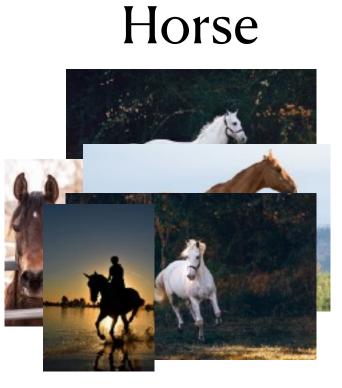


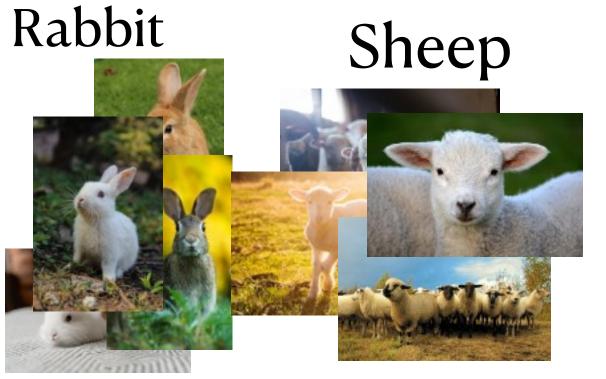


Dog



Fish





$$L \sim \mathcal{T}$$

 $S \sim L$

 $Q \sim L$

 $f_{\varphi}(image)$

Embedding Module

$$g_{\varphi}(e_1, e_2)$$

Relation Module

$$\varphi, \phi \leftarrow \arg\min_{\varphi, \phi} \mathbb{E}_{L \sim \mathcal{T}} \left[\mathbb{E}_{S \sim L, Q \sim L} \left[\mathcal{L}_{episode}(S, Q) \right] \right]$$

$$\mathcal{L}_{episode}(S, Q) = \sum_{i=1}^{m} \sum_{j=1}^{n} \left(r_{i,j} - 1(y_i = y_j) \right)^2$$