

Episode Based Training (2)



We are going to treat this as mini-batch training

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

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

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



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





Label of image i

Label of image j

Relation score of
query image j with
sample image i

Episode Based Training (2)

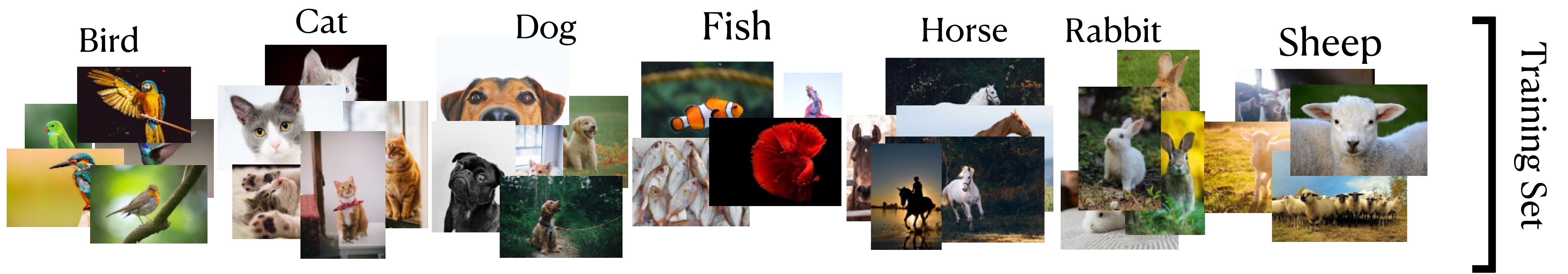
We are going to treat episodes as mini-batches in training

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

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

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

sampling a label set over a distribution over possible label sets \mathcal{T}



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

$$S \sim L$$

$$Q \sim L$$

$$f_{\phi}(image)$$

Embedding
Module

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

Relation
Module

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

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