

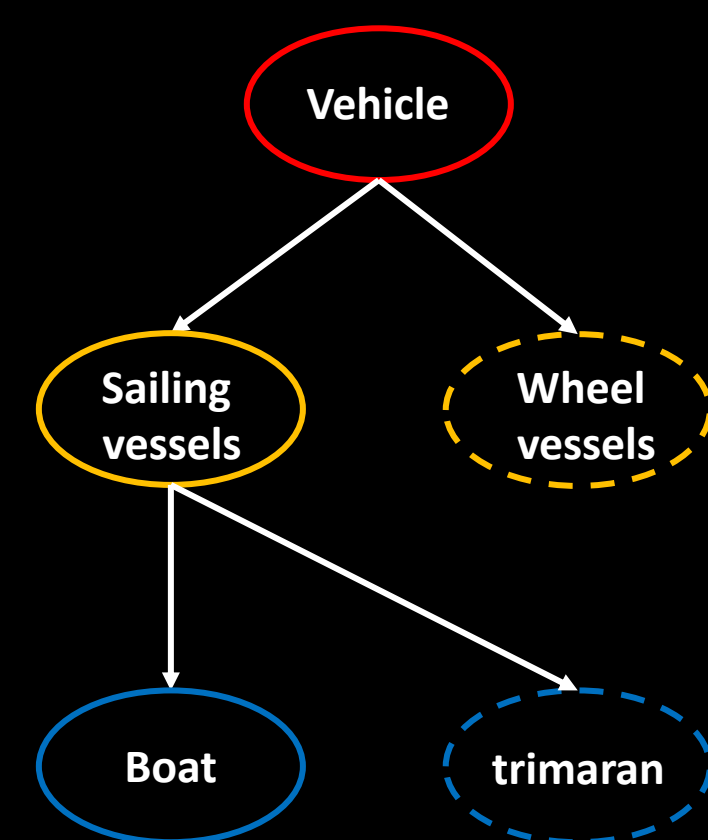
Relative Contrastive Loss for Self-supervised Learning

Shixiang Tang^{1,2}, Feng Zhu², Lei Bai^{1,3}, Rui Zhao², Wanli Ouyang^{1,3}

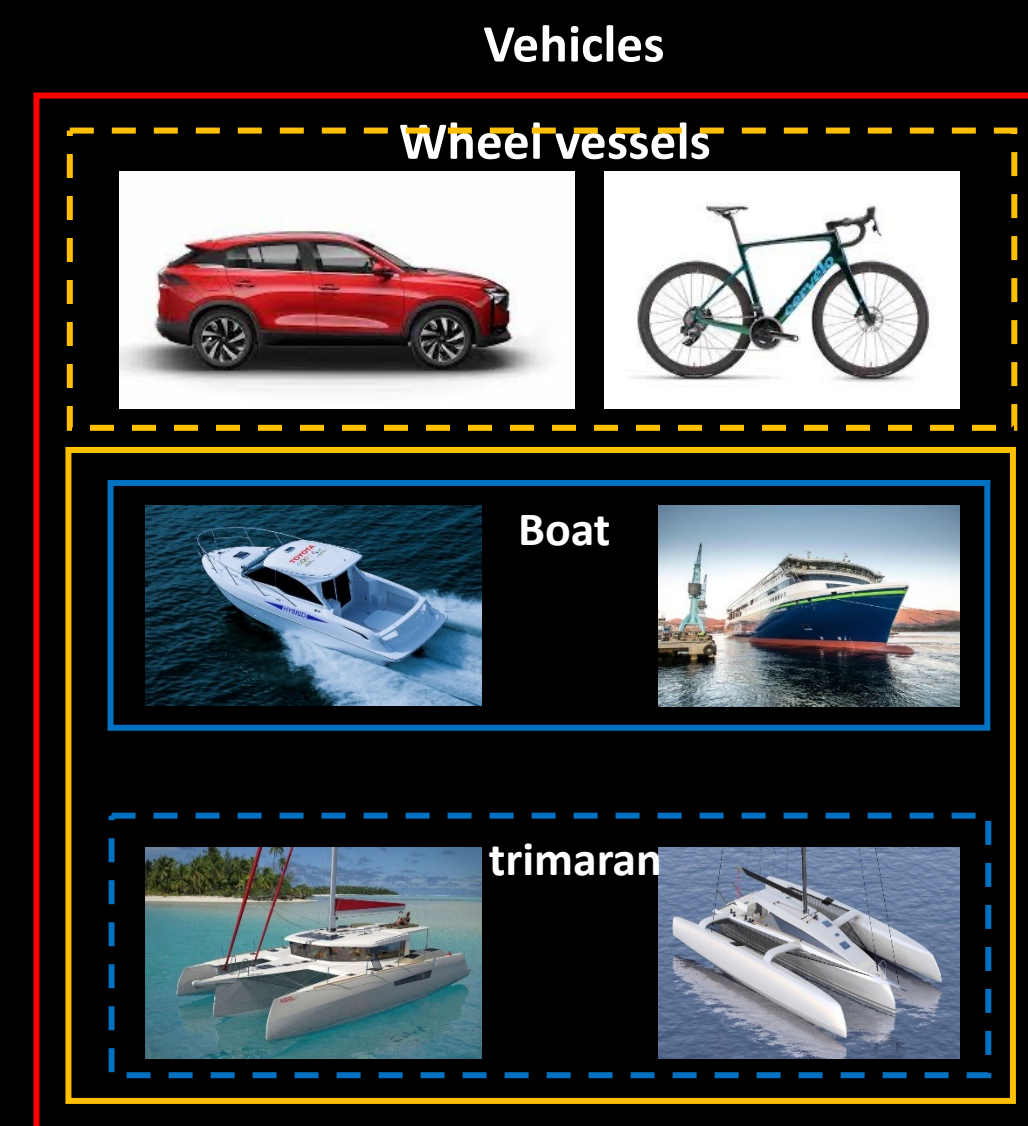
1 The University of Sydney, 2 SenseTime Group Limited, 3 Shanghai AI Laboratory

Motivation

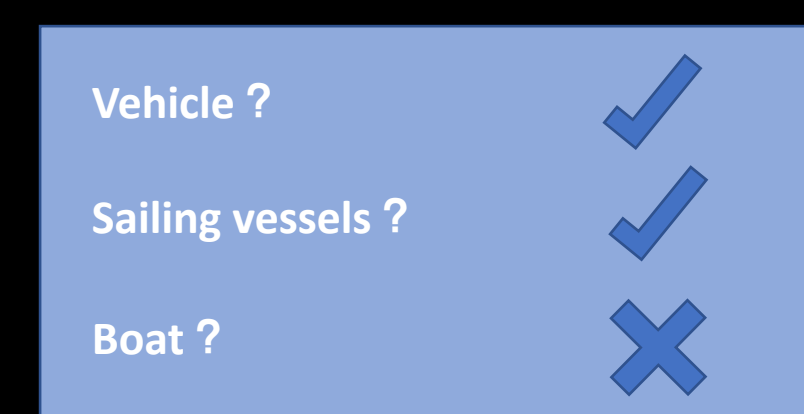
ImageNet: hierarchical organization



How to capture
hierarchical organizations
in ImageNet?

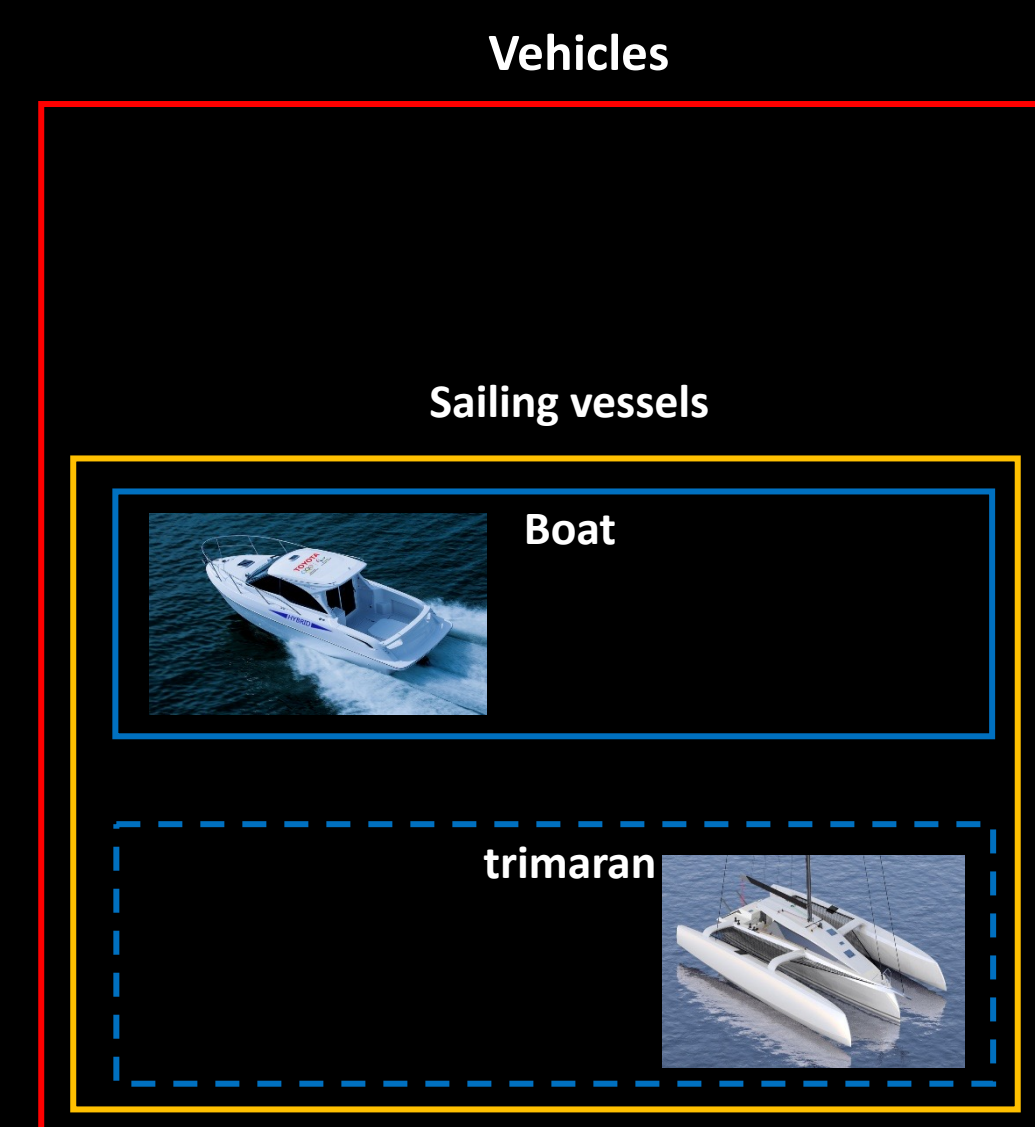


- Answer: Define relative positive and relative negative samples.



66.6% Positive

33.3% Negative



Relative Contrastive Loss

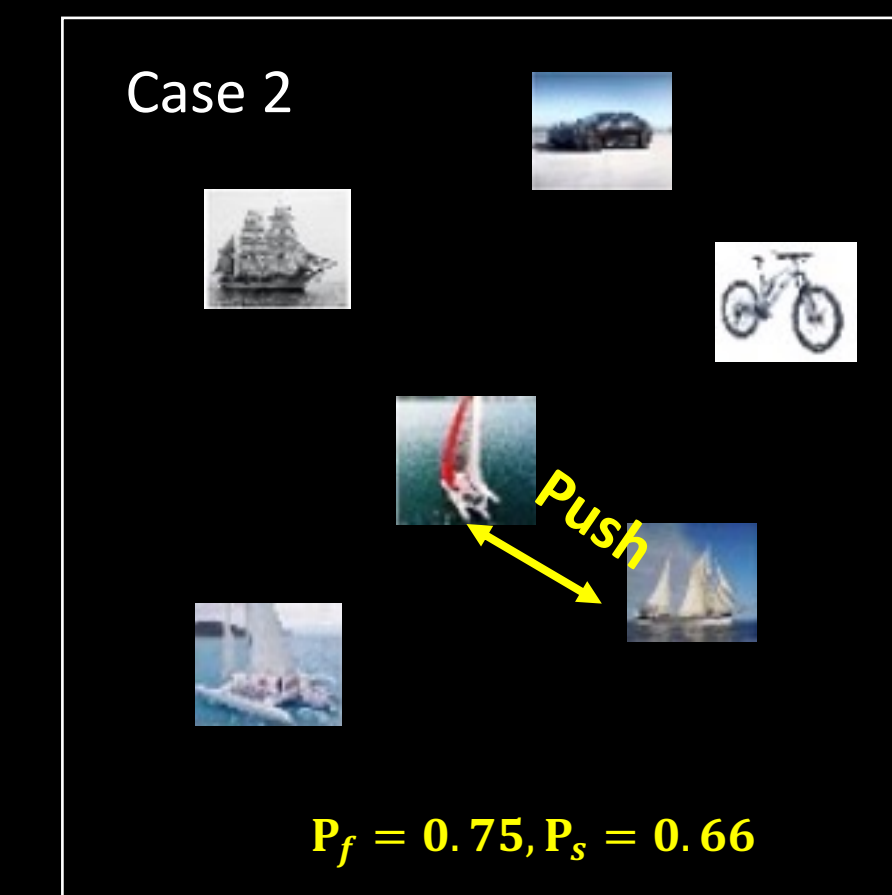
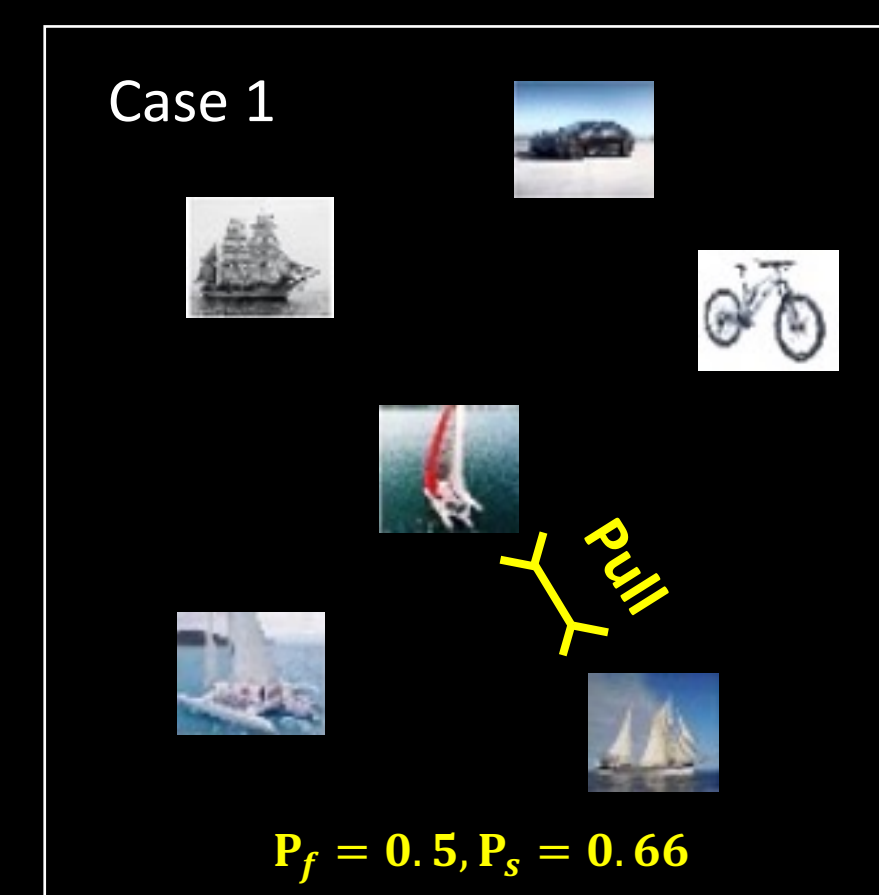
Give a set of criteria $\mathcal{M} = \{\mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_H\}$ to label a sample \mathbf{x} , i.e., $\mathcal{Y}(\mathbf{x}) = \{\mathcal{Y}_1(\mathbf{x}), \mathcal{Y}_2(\mathbf{x}), \dots, \mathcal{Y}_H(\mathbf{x})\}$,

The relative contrastive loss is defined by

$$\mathcal{L}_{RCL}(\mathbf{x}, \mathbf{x}'; \{\mathcal{Y}_i(\mathbf{x})\}_{i=1}^H, \{\mathcal{Y}_i(\mathbf{x}')\}_{i=1}^H) = \sum_{i=1}^H \mathcal{L}(\mathbf{x}, \mathbf{x}'; \mathcal{Y}_i(\mathbf{x}), \mathcal{Y}_i(\mathbf{x}'))$$

Typical Contrastive loss

$\mathcal{Y}_i(\mathbf{x}) = \mathcal{Y}_i(\mathbf{x}')$, pull;
 $\mathcal{Y}_i(\mathbf{x}) \neq \mathcal{Y}_i(\mathbf{x}')$, push;



visual feature similarity $P_f <$ positive-negative relativeness P_s , pull;

visual feature similarity $P_f >$ positive-negative relativeness P_s , push;

Pipeline

