



॥ नन्दं कर्मणो विद्यायमर्पयति ॥

Few-shot Visual Relationship Co-localization



Revant Teotia*, Vaibhav Mishra*, Mayank Maheshwari*, Anand Mishra

IIT Jodhpur, India

(*: equal contribution)

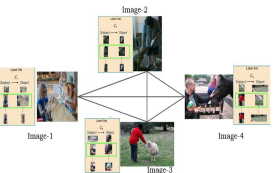


1 GOAL



Visual Relationship Co-Localization (VRC) in a bag of Images

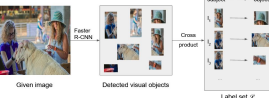
2 VRC AS LABELLING PROB.



Construct a fully connected graph:
Nodes: Images
Label set: all possible visual relationships

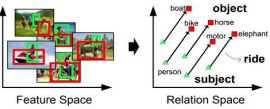
3 OUR APPROACH

Generating Label set for Images



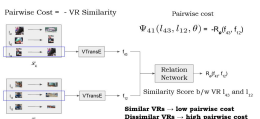
Faster-RCNN → Object proposals in images
 Each pair of object proposal = potential common visual relationship

Encoding Visual Relationship



Object proposal pair → VTransE embedding in visual relationship space

Learning Similarity with Meta-Learning Framework



Episodic training of Relation Network to predict visual relationship similarity

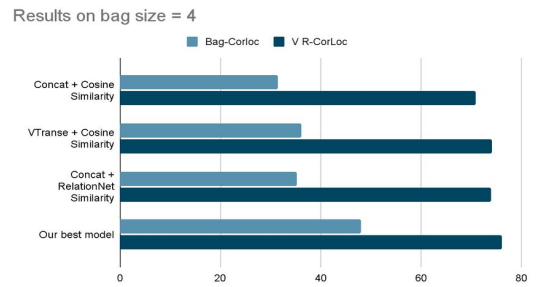
Greedy Inference

Large bag of images → smaller bag of images
 Solutions of smaller bags → combined greedily to get final solution in polynomial time.

4 RESULTS



Latent visual relationship: **sniffing**



5 SUMMARY

- A novel task of visual relationship co-localization
- A principled meta learning based optimization framework
- Opens up many future research avenues