

# Contrastive Learning for Weakly Supervised Phrase Grounding

(Supplementary Material)

Paper ID 1327

- To give the readers an accurate qualitative sense of performance, we include 20 randomly sampled results for our full model trained using COCO training data on Flickr30K Entities validation set.
- These results are **not** cherry-picked.
- For every image-caption pair, we show top ranking bounding boxes with scores for two words in the caption. The two words are nouns chosen from the sentence without looking at grounding results while excluding “stuff” (e.g. sky, snow) or “scene” (playground, sidewalk) categories.
- First 13 pairs are correct, remaining 7 are incorrect
- Note that in the partially correct or incorrect cases true regions are often ranked in the top 3 predictions if not the top prediction.

Correct Predictions

An old man is sitting on a sidewalk in front of a pile of folded blankets.



**#boxes**  
**26**



**man**  
**0.52, 0.15, 0.09**



**blankets**  
**0.29, 0.23, 0.12**



A child in a helmet plays on a skateboard by a playground.



**#boxes**

**34**



**helmet**

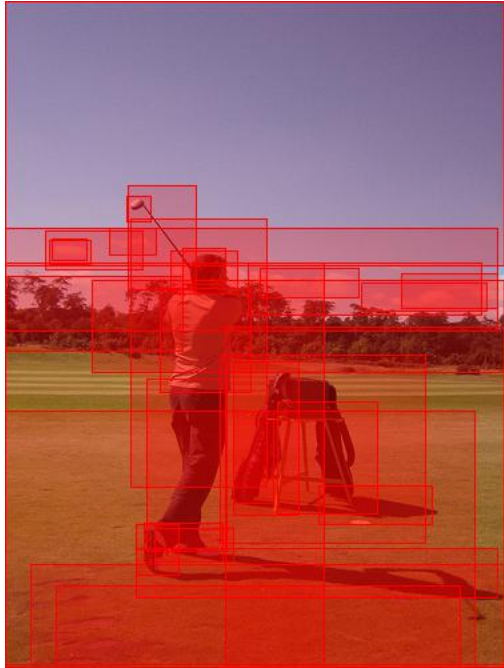
**0.57, 0.11, 0.04**



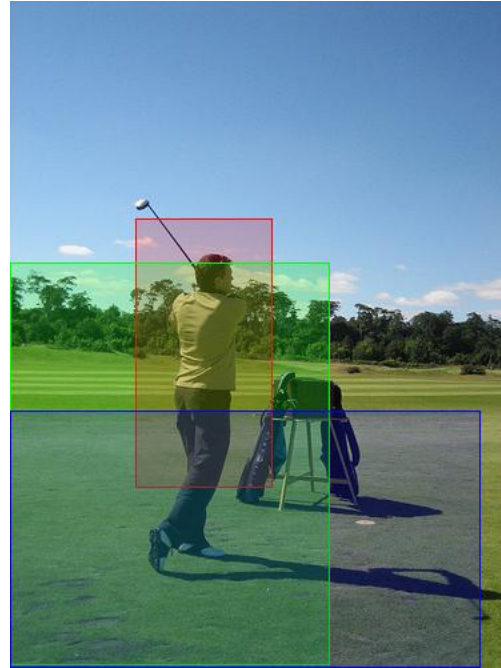
**skateboard**

**0.17, 0.09, 0.08**

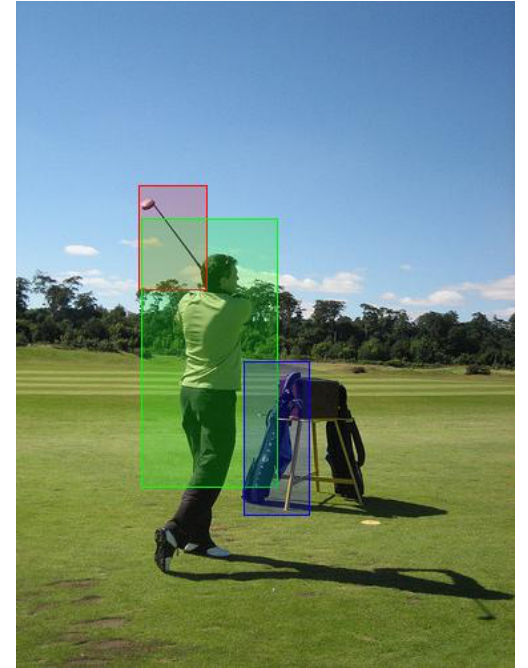
**Man** swinging a golf **club** at a golf course.



**#boxes**  
**33**



**man**  
**0.41, 0.12, 0.1**

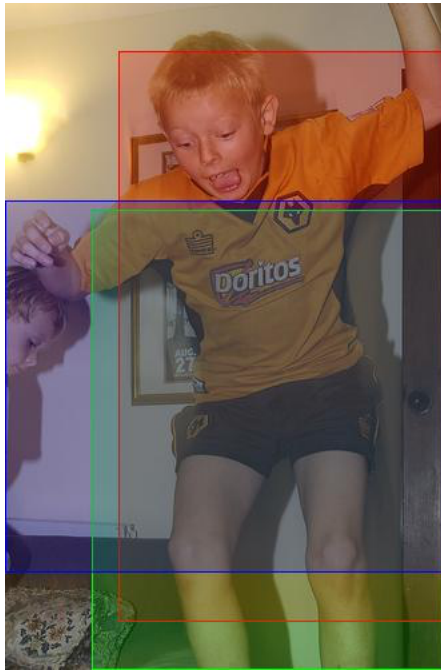


**club**  
**0.33, 0.17, 0.1**

A boy wearing an orange doritos shirt looks like he 's about to jump off of a piece of furniture.



**#boxes**  
**28**



**boy**  
**0.28, 0.11, 0.1**



**doritos**  
**0.27, 0.19, 0.11**



A little girl<sup>1</sup> with a yellow shirt swings while a little girl<sup>2</sup> in green watches.



**#boxes**  
**21**

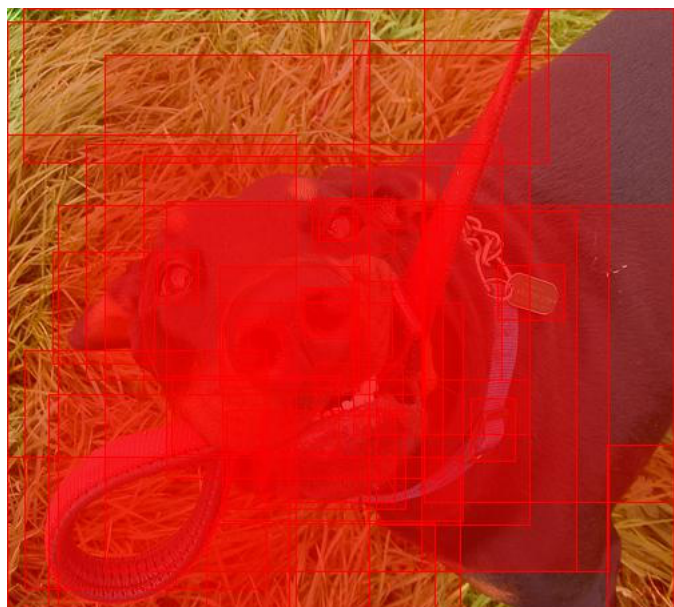


**girl<sup>1</sup>**  
**0.29, 0.12, 0.12**



**girl<sup>2</sup>**  
**0.2, 0.16, 0.11**

A black dog chewing on a red leash.



**#boxes**  
**35**



**dog**  
**0.35, 0.26, 0.11**

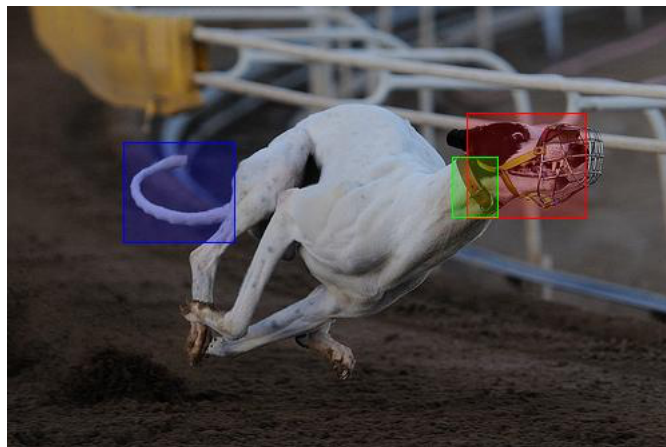


**leash**  
**0.24, 0.15, 0.14**

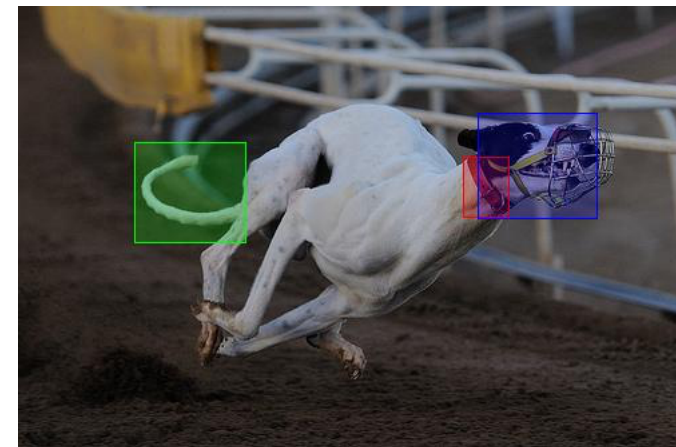
A white dog with a mouthpiece and a red collar running.



#boxes  
17



mouthpiece  
0.34, 0.21, 0.11



collar  
0.21, 0.19, 0.11



A guy waxing a wooden floor with wax.



**#boxes**  
**19**

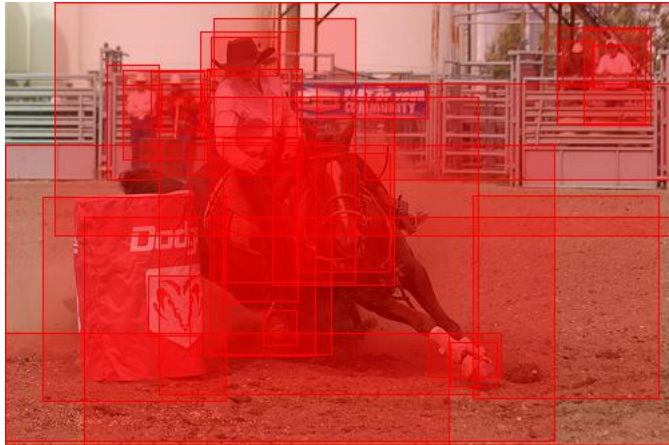


**guy**  
**0.32, 0.18, 0.08**

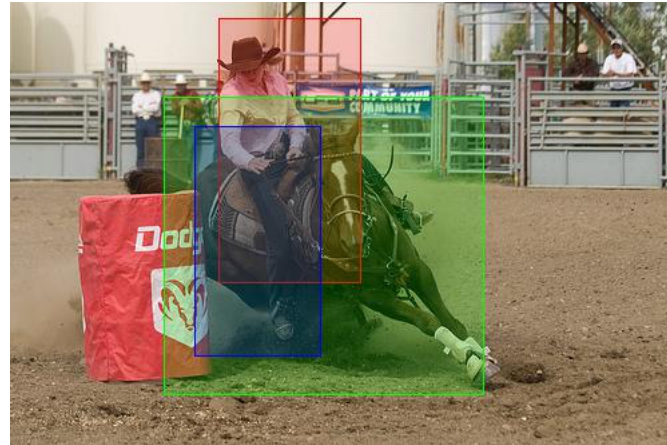


**wax**  
**0.33, 0.18, 0.15**

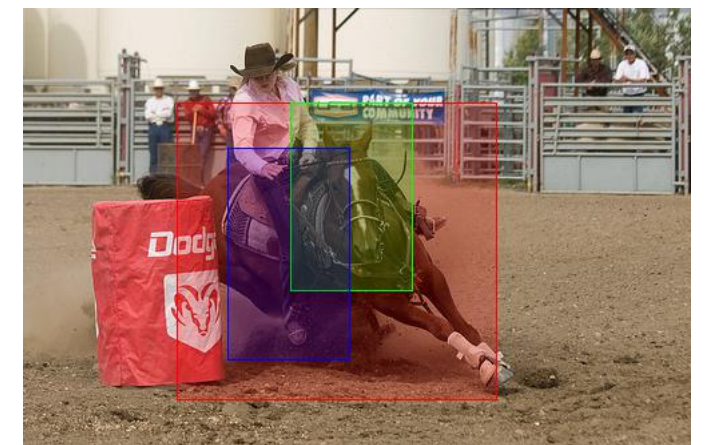
A woman in a pink shirt wears rides a horse in a rodeo.



**#boxes**  
**31**



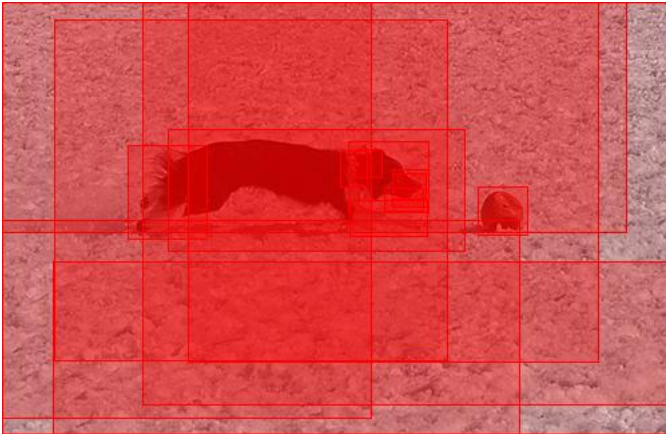
**woman**  
**0.27, 0.16, 0.08**



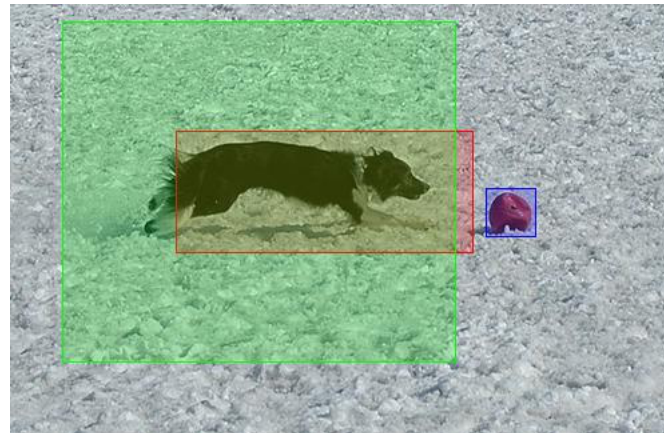
**horse**  
**0.36, 0.11, 0.11**



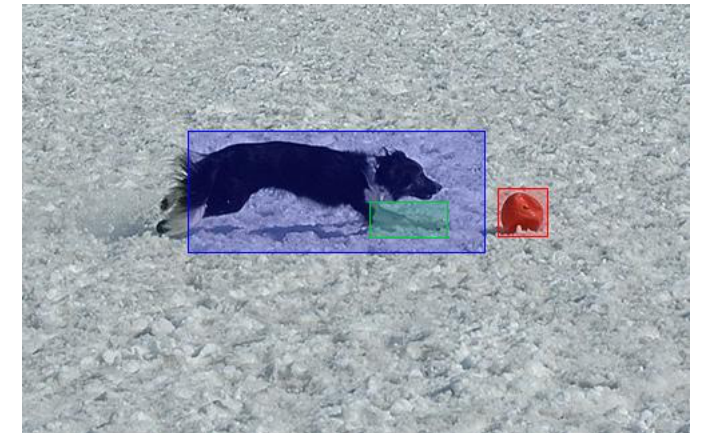
A black dog is chasing a punctured ball in the snow.



**#boxes**  
**17**



**dog**  
**0.46, 0.15, 0.08**



**ball**  
**0.82, 0.04, 0.02**

An Asian man wearing a blue long-sleeved shirt is overlooking another man that is butchering a fish.



#boxes  
24



shirt  
0.26, 0.24, 0.09



fish  
0.13, 0.12, 0.12

A little boy runs through a crowd of pigeons.



#boxes  
44



boy  
**0.53**, 0.15, 0.08



pigeons  
**0.07**, 0.07, 0.07



A group of people playing a board game.



#boxes  
37



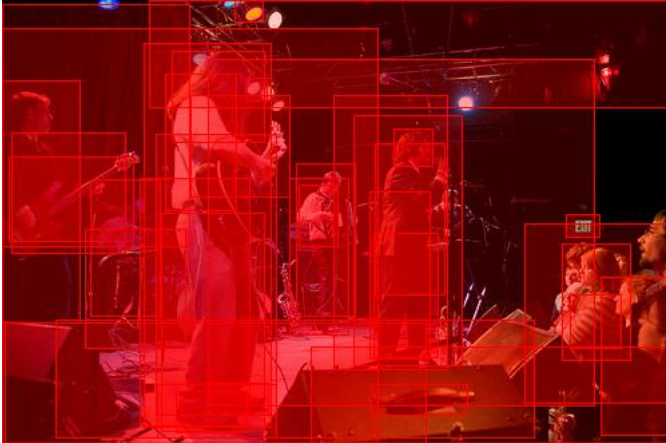
people  
0.2, 0.11, 0.09



board  
0.15, 0.08, 0.08

Partially Correct or Incorrect  
Predictions

A band performs music for a crowd of people.



#boxes  
39

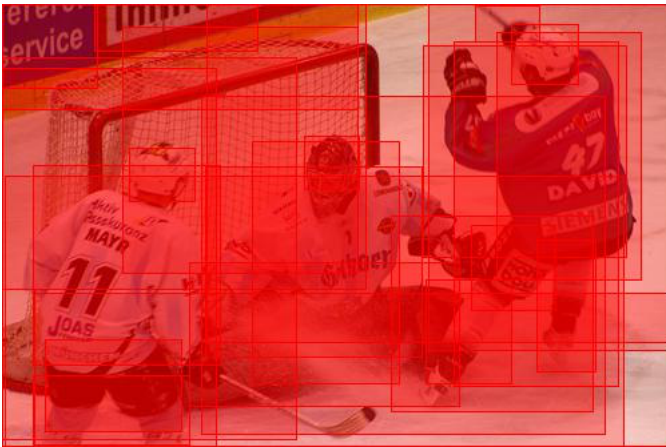


band  
0.17, 0.1, 0.09



crowd  
0.13, 0.06, 0.06

A hockey player attempting to score against a goalie as a defenceman watches.



#boxes  
29



player  
0.12, 0.12, 0.09



goalie  
0.11, 0.1, 0.08



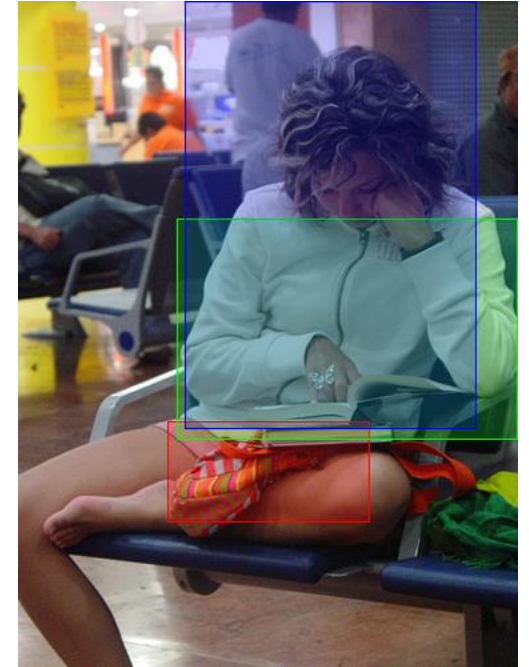
A woman in a white **jacket** and pink **shorts** is sitting down, reading a book with her head resting on her right hand.



**#boxes**  
**32**



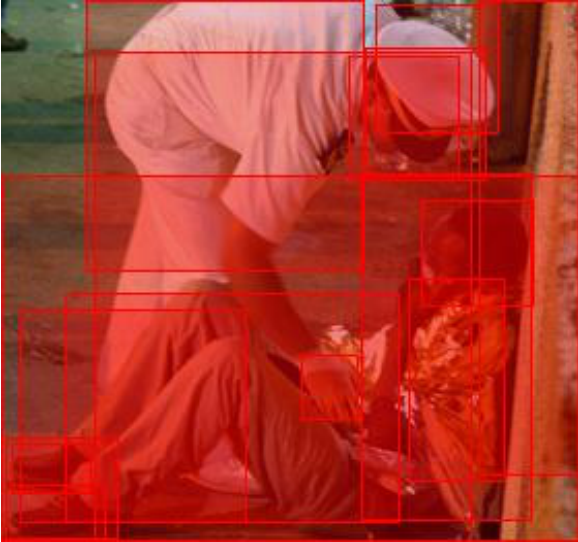
**jacket**  
**0.17, 0.17, 0.11**



**shorts**  
**0.23, 0.09, 0.07**



A man<sup>1</sup> sitting on the floor crying while another man<sup>2</sup> is trying to comfort him.



**#boxes**  
**16**



**man<sup>1</sup>**  
**0.38, 0.21, 0.2**



**man<sup>2</sup>**  
**0.25, 0.21, 0.19**

The artist paints a picture of a beautiful building during the winter.



#boxes  
28

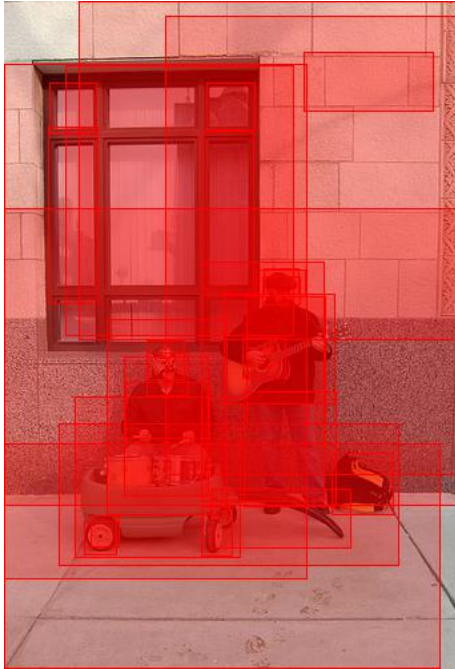


artist  
0.16, 0.14, 0.07

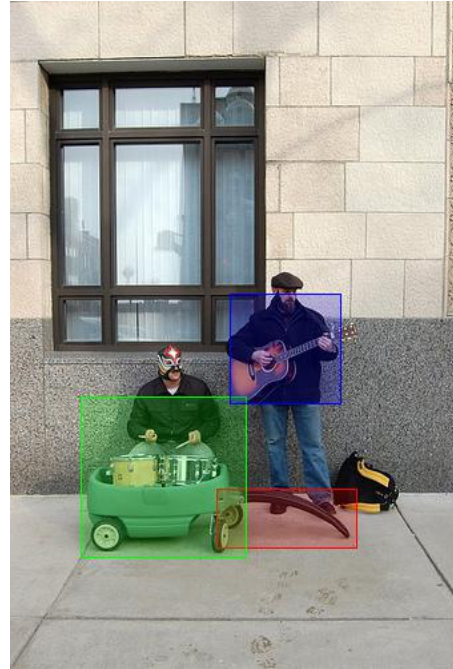


building  
0.15, 0.09, 0.07

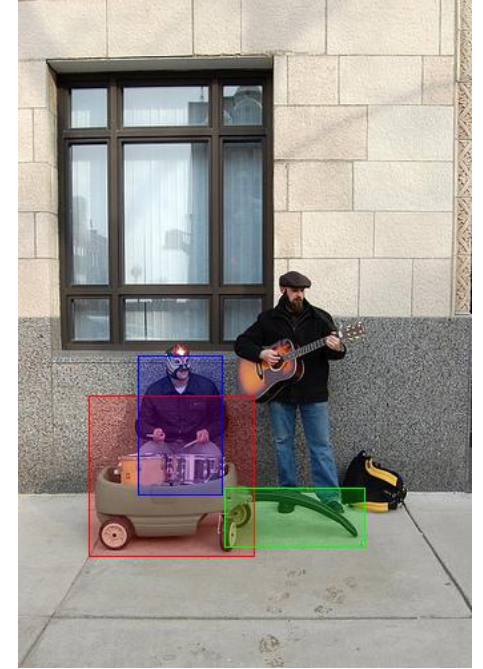
Two men next to a building playing a guitar and drums.



#boxes  
31

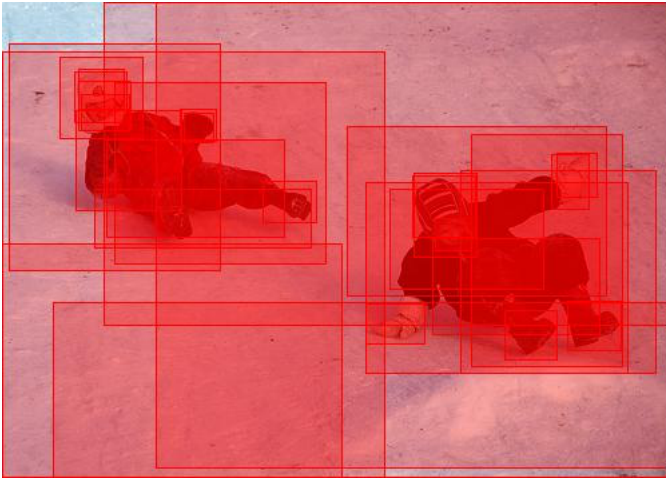


guitar  
0.19, 0.11, 0.09



drums  
0.19, 0.11, 0.08

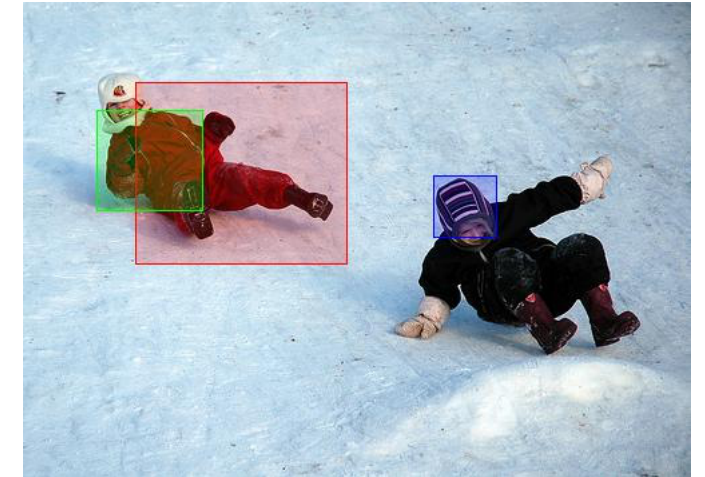
Two small children, one wearing a red winter jacket and the other a black one, are slipping and sliding down an icy slope.



**#boxes**  
**32**



**red**  
**0.11, 0.1, 0.1**



**black**  
**0.1, 0.07, 0.07**



Thank You!