# A Distributed Representation Based Query Expansion Approach for Image Captioning (Supplementary Material)

**Semih Yagcioglu**[1]   **Erkut Erdem**[1]   **Aykut Erdem**[1]   **Ruket Çakıcı**[2]

[1] Hacettepe University Computer Vision Lab (HUCVL)

Dept. of Computer Engineering, Hacettepe University, Ankara, TURKEY

semih.yagcioglu@hacettepe.edu.tr, {erkut,aykut}@cs.hacettepe.edu.tr

[2] Dept. of Computer Engineering, Middle East Technical University, Ankara, TURKEY

ruken@ceng.metu.edu.tr

## 1 Additional Results

In Figure 1, we present more example results obtained with our approach on the benchmark datasets Flickr8K (Hodosh et al., 2013), Flickr30K (Young et al., 2014), MS COCO (Lin et al., 2014). We also provide groundtruth human descriptions for comparison.

There are some cases where our approach falls short. In some of those cases, although the system does not produce the most desirable results, it often is able to produce interesting results as it could capture some of the semantic relations correctly. We provide such poetic or funny results in Figure 2.

There are a few cases where our system could not generate desired outputs. This is generally due to the fact that the retrieval set does not contain representative captions either because the system fails to retrieve visually similar matches or there is more noise in the retrieved captions that our system could handle. We also provide such results in Figure 3.

## References

Jonathan H Clark, Chris Dyer, Alon Lavie, and Noah A Smith. 2011. Better hypothesis testing for statistical machine translation: Controlling for optimizer instability. In *Proc. of ACL*.

Micah Hodosh, Peter Young, and Julia Hockenmaier. 2013. Framing Image Description as a Ranking Task: Data, Models and Evaluation Metrics. *Journal of Artificial Intelligence Research*. 1

Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. 2014. Microsoft COCO: Common Objects in Context. In *Proc. of ECCV*. 1

Rebecca Mason and Eugene Charniak. 2014. Nonparametric Method for Data-driven Image Captioning. In *Proc. of ACL*.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed Representations of Words and Phrases and their Compositionality. In *Proc. of NIPS*.

Jeffrey Pennington, Richard Socher, and Christopher D Manning. 2014. GloVe: Global Vectors for Word Representation. *Proc. of EMNLP*.

Peter Young, Alice Lai, Micah Hodosh, and Julia Hockenmaier. 2014. From Image Descriptions to Visual Denotations: New similarity Metrics for Semantic Inference over Event Descriptions. *Transactions of the Association for Computational Linguistics*. 1

|  | OURS | HUMAN |
|---|---|---|
|  | a black dog and a brown dog touching noses in a grassy field | two dogs are fighting over something |
|  | a tan dog is running through shallow water with a stick in its mouth | a dog is fetching a stick from a lake |
|  | a skateboarder in midair above an indoor ramp | a lone skateboarder jumping a huge skateboarding ramp |
|  | some musicians dressed in native american costume are playing traditional instruments | a man dressed as an indian speaking into a microphone |
|  | two girls in pink are playing on yellow playground bars | the two girls are playing on a yellow and red jungle gym |
|  | a brown dog is running through the snow | a dog in the snow |
|  | a person wearing a red jacket climbs a snowy hill | a hiker in a red on a snowy peak |

Figure 1: **Additional Results.** Input images on the left, our results in the middle and groundtruth human descriptions on the right.
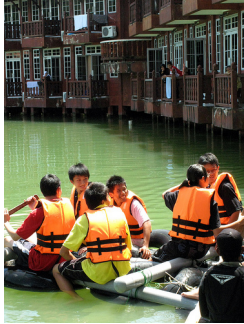
|  | OURS | HUMAN |
|---|---|---|
|  | two young children play with water in a fountain | two girls running through pool while water sprays over them |
|  | boys in orange life jackets sitting on raft in water | a group of six boys are wearing yellow life vests and are on a make-shift raft |
|  | a young boy on a basketball team dribbles the ball | a boy playing basketball in a gym |
|  | little boy in a red jacket sliding in the snow | a child dressed for the cold sits in the snow |
|  | two people in a canoe on a river beside a forested mountain | a kayak on the river is passing by a wire bridge that joins both banks |
|  | a construction crew in orange vests working near train tracks | a man in an orange vest and yellow hard hat looks on as a yellow vehicle lays track |

Figure 1: (continued) **Additional Results.** Input images on the left, our results in the middle and groundtruth human descriptions on the right.

|  | OURS | HUMAN |
|---|---|---|
|  | woman in blue is tossing a young child into a pool | a boy wearing red shorts is jumping into a paddling pool |
|  | two brown dogs race on a track | the greyhounds are running quickly in this race |
|  | a man skis on a snowy mountain | a man rides a ski lift up a mountain |
|  | a close-up of a young girl with a pink shirt laying on grass | a girl making a sad face standing in front of her tricycle |
|  | a green bird perched on top of a tree filled with pink flowers | grey plane taking off above some green vegetation |
|  | a white cat is sitting in a bathroom sink | a small black cat sitting on the ground |
|  | a brown dog runs for a white and black dog on the grass | a dog biting the tail of a cow |
|  | a kitchen with a white sink and a stove top oven | an oven and a range on a table |

Figure 2: **Funny Results.** Input images on the left, our results in the middle and groundtruth human descriptions on the right.

|  | OURS | HUMAN |
|---|---|---|
|  | three people are standing on rocks near a body of water | a man stands alone on the rocks |
|  | a wooden table topped with a bowl of soup and pizza | fast food displayed on a table with sandwich and soup |
|  | oranges and apples for sale in a market | a grocery store display filled with lots of apples |
|  | a man in a blue jacket with a helmet on skiing | a female skier in a green with pink flowers hat skiing |
|  | a blue double decker city bus driving down a street | a very big city bus on a big street |
|  | a man and a woman are riding bikes on a street | an elderly woman is riding a bicycle in the city as a yellow taxi is about to pass by |
|  | a smiling man wearing a red hat and blue jacket | a female police officer wearing an officers hat and sunglasses stands in uniform in front of a window-lined street block |

Figure 3: **Undesired Results.** Input images on the left, our results in the middle and groundtruth human descriptions on the right.