5th Workshop on Vision and Language (VL’16)

Held at ACL 2016

Berlin, Germany
12 August 2016

ISBN: 978-1-5108-2763-9

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571

Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2016) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2016)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

Automatic Annotation of Structured Facts in Images
Mohamed Elhoseiny, Scott Cohen, Walter Chang, Brian Price and Ahmed Elgammal 1

Combining Lexical and Spatial Knowledge to Predict Spatial Relations between Objects in Images
Manuela Hürlimann and Johan Bos 10

Focused Evaluation for Image Description with Binary Forced-Choice Tasks
Micah Hodosh and Julia Hockenmaier 19

Leveraging Captions in the Wild to Improve Object Detection
Mert Kilickaya, Nazli Ikizler-Cinbis, Erkut Erdem and Aykut Erdem 29

Natural Language Descriptions of Human Activities Scenes: Corpus Generation and Analysis
Nouf Alharbi and Yoshihiko Gotoh 39

Interactively Learning Visually Grounded Word Meanings from a Human Tutor
Yanchao Yu, Arash Eshghi and Oliver Lemon 48

Pragmatic Factors in Image Description: The Case of Negations
Emiel van Miltenburg, Roser Morante and Desmond Elliott 54

Building a Bagpipe with a Bag and a Pipe: Exploring Conceptual Combination in Vision
Sandro Pezzelle, Ravi Shekhar and Raffaella Bernardi 60

Exploring Different Preposition Sets, Models and Feature Sets in Automatic Generation of Spatial Image Descriptions
Anja Belz, Adrian Muscat and Brandon Birmingham 65

Multi30K: Multilingual English-German Image Descriptions
Desmond Elliott, Stella Frank, Khalil Sima'an and Lucia Specia 70

"Look, some Green Circles!": Learning to Quantify from Images
Ionut Sorodoc, Angeliki Lazaridou, Gemma Boleda, Aurélie Herbelot, Sandro Pezzelle and Raffaella Bernardi 75

Text2voronoi: An Image-driven Approach to Differential Diagnosis
Alexander Mehler, Tolga Uslu and Wahed Hemati 80

Detecting Visually Relevant Sentences for Fine-Grained Classification
Olivia Winn, Madhavan Kavanur Kidambi and Smaranda Muresan 86