

NAACL HLT 2019

**Second Workshop on Shortcomings in Vision and Language
(SiVL)**

Proceedings of the Workshop

June 6, 2019
Minneapolis, USA



©2019 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-950737-07-9

Introduction

Welcome to the second edition of Workshop on Shortcomings in Vision and Language (SiVL) at NAACL. The first installment of SiVL was held at ECCV 2018 (the European Conference on Computer Vision). SiVL’s primary purpose is to bring together researchers at the intersection of vision and language to discuss shortcomings of modern approaches, tasks, data-sets, and evaluation metrics. By highlighting common shortcomings in these domains, the workshop aims to facilitate discussion of novel research directions and to steer the community towards high-level challenges affecting the vision and language community broadly. This year, we called both for Full Papers and Extended Abstracts. The Full Papers are published in these proceedings whereas the extended abstracts are available from the workshop website.

The works in these proceedings highlight the following shortcomings of current tasks, data-sets and metrics.

One main important issue concerns the difference between data-sets annotated by subjects on demand versus naturalistic ones. Because of these differences, models trained on the former are not necessarily suitable for the latter. Findings from some of the submitted works call for new and richer data-sets that exhibit semantic and pragmatic diversity as well as new metrics that take these qualitative aspects of the generated data into account.

Another important concern is about bias in visual and language grounding tasks. Works studying the bias problem points at the importance of finding models that can cope with biases, maybe even exploit them when data are aligned with bias, but override them when they are not. Crucially, in-depth evaluation analyses have to be conducted to guarantee the stability of the models’ results.

SiVL-2019 received 13 valid full-paper submissions among which 2 were withdrawn before reviewing phase. The remaining papers received a minimum of 3 double-blind reviews from our highly qualified program committee. Among these, we accepted 8 papers to appear in the workshop. We also received 7 extended abstract submissions, all of which were accepted to appear in the workshop based on the topic relevance determined by the organizing committee.

In addition to full papers and extended abstracts, SiVL-2019 program also includes three invited talks by Yoav Artzi, Angeliki Lazaridou and Margaret Mitchell, thanks to our sponsor SAP. SiVL has also received a sponsorship by Google AI Language.

Organizers:

Raffaella Bernardi, University of Trento (Italy)
Raquel Fernandez, University of Amsterdam (Netherlands)
Spandana Gella, Amazon AI (USA)
Kushal Kafle, Rochester Institute of Technology (USA)
Christopher Kanan, Rochester Institute of Technology (USA)
Stefan Lee, Georgia Tech (USA)
Moin Nabi, SAP SE (Germany)

Program Committee:

Aaditya Prakash, Brandeis University
Abhisek Das, Georgia Tech
Aishwarya Agrawal, Georgia Tech
Alane Suhr, Cornell University
Albert Gatt, University of Malta
Anna Rohrbach, University of California, Berkeley
Aurelie Herbelot, University of Trento
Brandon Birmingham, University of Malta
Carina Silberer, Universitat Pompeu Fabra
Damien Teney, The University of Adelaide
David Schlangen, Bielefeld University
Desmond Elliott, University of Copenhagen
Douwe Kiela, Facebook AI Research
Elia Bruni, University of Amsterdam
Emiel van Miltenburg, Tilburg University
Enver Sangineto, University of Trento
Iacer Calixto, University of Amsterdam
Ivan Vulic, University of Cambridge
Jasper Uijlings, Google AI
Justin Johnson, Stanford University
Katerina Pastra, Cognitive Systems Research Institute
Lisa Anne Hendricks, UC Berkeley
Lucia Specia, University of Sheffield
Malihe Alikhani, Rutgers University
Manoj Acharya, Rochester Institute of Technology
Nazli Ikizler-Cinbis, Hacettepe University
Peratham Wiriathamabhum, University of Maryland, College Park
Pranava Madhyastha, University of Sheffield
Ramakanth Pasunuru, UNC Chapel Hill
Ranjay Krishna, Stanford University
Robik Shreshtha, Rochester Institute of Technology
Sandro Pezzelle, University of Amsterdam
Sina Zarriess, Bielefeld University
Volkan Cirik, Carnegie Mellon University
Yash Goyal, Georgia Tech

Invited Speaker:

Yoav Artzi, Cornell Tech

Angeliki Lazaridou, Deepmind

Margaret Mitchell, Google Research

Table of Contents

<i>Adversarial Regularization for Visual Question Answering: Strengths, Shortcomings, and Side Effects</i> Gabriel Grand and Yonatan Belinkov	1
<i>Referring to Objects in Videos Using Spatio-Temporal Identifying Descriptions</i> Peratham Wiriyathammabhum, Abhinav Shrivastava, Vlad Morariu and Larry Davis	14
<i>A Survey on Biomedical Image Captioning</i> John Pavlopoulos, Vasiliki Kougia and Ion Androutsopoulos	26
<i>Revisiting Visual Grounding</i> Erik Conser, Kennedy Hahn, Chandler Watson and Melanie Mitchell	37
<i>The Steep Road to Happily Ever after: an Analysis of Current Visual Storytelling Models</i> Yatri Modi and Natalie Parde	47
<i>"Caption" as a Coherence Relation: Evidence and Implications</i> Malihe Alikhani and Matthew Stone	58
<i>Learning Multilingual Word Embeddings Using Image-Text Data</i> Karan Singhal, Karthik Raman and Balder ten Cate	68
<i>Grounded Word Sense Translation</i> Chiraag Lala, Pranava Madhyastha and Lucia Specia	78

Conference Program

Shortcomings in vision and language (SiVL) is a full-day event with three plenary talks, spotlight talks for accepted full paper submissions and poster presentations for all accepted full papers and extended abstracts. Presenters are welcome to present their posters during both sessions, but at the minimum are expected to be present during their designated session, i.e., session 1 for abstracts and session 2 for full papers.

The full schedule is presented below.

Thursday, June 6, 2019

09:00–09:10 *Welcome and Opening Remarks*

09:10–10:00 *Invited Talk 1: Angeliki Lazaridou*

10:00–11:20 *Poster Session 1: Abstracts (with coffee break 10:30 - 11:00)*

11:20–12:30 *Spotlight Talks (Full Papers)*

12:30–14:10 *Lunch*

14:10–14:50 *Invited Talk 2: Margaret Mitchell*

14:50–16:10 *Poster Session 2: Full Papers (with coffee break 15:30 - 16:00)*

16:10–17:00 *Invited Talk 3: Yoav Artzi*

17:00–17:30 *Closing Remarks and Panel Discussion*

