

EMNLP-IJCNLP 2019

**Beyond Vision and LANguage: inTEgrating Real-world kNowledge
(LANTERN)**

Proceedings of the Workshop

November 3, 2019
Hong Kong, China



©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 978-1-950737-75-8

Introduction

Welcome to The First Workshop Beyond Vision and Language: Integrating Real-World Knowledge (LANTERN) co-located with EMNLP-IJCNLP 2019. The primary goal of the workshop is to bring together researchers adopting machine learning techniques to interconnect language, vision, and other modalities by leveraging external knowledge. By encouraging contributions exploiting very diverse sources of external knowledge (knowledge graphs, fixed and dynamic environments, cognitive and neuroscience data, etc.), the workshop aims to foster discussion and promote novel research directions which acknowledge the importance of knowledge in acquiring, using, and evaluating language in real-world settings.

In this first edition, we called for both long and short papers. All the accepted ones are published in these Proceedings.

LANTERN 2019 received 17 submissions, out of which 1 was desk-rejected before the reviewing phase due to inappropriateness. The remaining 16 papers received 2 highly-qualified double-blind reviews. Besides considering the average overall score, only papers for which none of the reviewers expressed a negative opinion (strong or weak reject) were accepted. In total, 9 papers (3 long, 6 short) were accepted to appear in the workshop, with an acceptance rate of around 53%.

Contributions are representative of a varied number of current problems and approaches and include novel deep learning techniques and tasks in the domain of visual reasoning, the use of multilingual embeddings and scene graphs in multimodal tasks, multi-agent language learning, the use of eye-tracking data in syntactic tagging and of textual adversaries in multimodal machine translation, the extraction of orthography knowledge from Chinese characters. Such richness of approaches and perspectives is in line with the purpose of the workshop, and confirms the growing interest for problems going beyond the task-specific integration of language and vision.

The program of the workshop, besides 4 oral presentations and a poster session, includes 3 invited talks by Lucia Specia, Mohit Bansal, and Siddharth Narayanaswamy. The workshop received sponsorship by iDeaL SFB 1102 and Nuance. Best paper and best poster awards are sponsored by SAP.

The LANTERN Workshop Organizers

Organizers:

Aditya Mogadala, Saarland University (Germany)
Dietrich Klakow, Saarland University (Germany)
Sandro Pezzelle, University of Amsterdam (The Netherlands)
Marie-Francine Moens, Katholieke Universiteit Leuven (Belgium)

Program Committee:

Spandana Gella, Amazon AI
Ravi Shekhar, Queen Mary University London
Angeliki Lazaridou, DeepMind
Ionut-Teodor Sorodoc, Universitat Pompeu Fabra
Niket Tandon, Allen Institute for Artificial Intelligence (AI2)
Florian Metze, CMU
Alexander Kuhnle, Cambridge University
Mirella Lapata, Edinburgh University
Mario Fritz, CISPA, Saarland
Carina Silberer, Universitat Pompeu Fabra
Barbara Plank, IT University of Copenhagen
Jacob Goldberger, Bar Ilan
Pranava Madhyastha, Imperial College London
Stefan Lee, Oregon State University
Parisa Kordjamshidi, Michigan State University

Invited Speaker:

Mohit Bansal, UNC Chapel Hill
Lucia Specia, Imperial College London
Siddharth Narayanaswamy, Oxford Univeristy

Table of Contents

<i>Structure Learning for Neural Module Networks</i> Vardaan Pahuja, Jie Fu, Sarath Chandar and Christopher Pal	1
<i>Aligning Multilingual Word Embeddings for Cross-Modal Retrieval Task</i> Alireza Mohammadshahi, Rémi Lebrete and Karl Aberer	11
<i>Big Generalizations with Small Data: Exploring the Role of Training Samples in Learning Adjectives of Size</i> Sandro Pezzelle and Raquel Fernández	18
<i>Eigencharacter: An Embedding of Chinese Character Orthography</i> Yu-Hsiang Tseng and Shu-Kai Hsieh	24
<i>On the Role of Scene Graphs in Image Captioning</i> Dalin Wang, Daniel Beck and Trevor Cohn	29
<i>Understanding the Effect of Textual Adversaries in Multimodal Machine Translation</i> Koel Dutta Chowdhury and Desmond Elliott	35
<i>Learning to request guidance in emergent language</i> Benjamin Kolb, Leon Lang, Henning Bartsch, Arwin Gansekoele, Raymond Koopmanschap, Leonardo Romor, David Speck, Mathijs Mul and Elia Bruni	41
<i>At a Glance: The Impact of Gaze Aggregation Views on Syntactic Tagging</i> Sigrid Klerke and Barbara Plank	51
<i>Seeded self-play for language learning</i> Abhinav Gupta, Ryan Lowe, Jakob Foerster, Douwe Kiela and Joelle Pineau	62

Conference Program

LANTERN is a half-day event with four oral and five spotlight presentations for the accepted long and short papers. It also includes three invited presentations. The full schedule is presented below.

14:00–14:15 Welcome and Opening Remarks

14:15–14:45 Invited Talk: Lucia Specia (Imperial College London)

Session 1: Accepted Papers (Oral)

14:45–15:00 On the Role of Scene Graphs in Image Captioning

15:00–15:15 Learning to request guidance in emergent language

15:15–15:30 Spotlight (posters)

- Aligning Multilingual Word Embeddings for Cross-Modal Retrieval Task
- Big Generalizations with Small Data: Exploring the Role of Training Samples in Learning Adjectives of Size
- Eigencharacter: An Embedding of Chinese Character Orthography
- Understanding the Effect of Textual Adversaries in Multimodal Machine Translation
- Seeded self-play for language learning

15:30–16:15 Coffee Break with Poster Session

16:15–16:45 Invited Talk: Siddharth Narayanaswamy (Oxford University)

Session 2: Accepted Papers (Oral)

16:45–17:00 At a Glance: The Impact of Gaze Aggregation Views on Syntactic Tagging

17:00–17:15 Structure Learning for Neural Module Networks

17:15–17:45 Invited Talk: Mohit Bansal (UNC Chapel Hill)

17:45–18:00 Best Talk/Poster Award and Concluding Remarks

