ACL 2020

The 58th Annual Meeting of the
Association for Computational Linguistics

Proceedings of the Second Grand Challenge and Workshop on Multimodal Language (Challenge-HML)

July 10, 2020
Online (due to COVID-19 pandemic)

©2020 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 978-1-952148-24-8 (Volume 1)

Introduction

The Second Grand-Challenge and Workshop on Multimodal Language (Challenge-HML) offers a unique opportunity for interdisciplinary researchers to study and model interactions between modalities of language, vision, and acoustic. This is the continuation of the Challenge-HML at ACL 2018. Modeling multimodal language is a growing research area in NLP. This research area pushes the boundaries of multimodal learning and requires advanced neural modeling of all three constituent modalities. Advances in this research area allow the field of NLP to take the leap towards better generalization to real-world communication (as opposed to limitation to textual applications), and better downstream performance in Conversational AI, Virtual Reality, Robotics, HCI, Healthcare, and Education.

Organizers:

Amir Zadeh, Language Technologies Institute, Carnegie Mellon University
Louis-Philippe Morency, Language Technologies Institute, Carnegie Mellon University
Paul Pu Liang, Machine Learning Department, Carnegie Mellon University
Soujanya Poria, Singapore University of Technology and Design

Invited Speakers:

Rada Mihalcea, University of Michigan (USA)
Ruslan Salakhutdinov, Carnegie Mellon University (USA)
M. Ehsan Hoque, University of Rochester (USA)
Yejin Choii, Washington University (USA)

Table of Contents

A Transformer-based joint-encoding for Emotion Recognition and Sentiment Analysis
Jean-Benoit Delbrouck, Noé Tits, Mathilde Brousmiche and Stéphane Dupont 1

A Multi-modal Approach to Fine-grained Opinion Mining on Video Reviews
Edison Marrese-Taylor, Cristian Rodriguez, Jorge Balazs, Stephen Gould and Yutaka Matsuo . . . 8

Multilogue-Net: A Context-Aware RNN for Multi-modal Emotion Detection and Sentiment Analysis in Conversation
Aman Shenoy and Ashish Sardana 19

Low Rank Fusion based Transformers for Multimodal Sequences
Saurav Sahay, Eda Okur, shachi H Kumar and Lama Nachman 29

Unsupervised Online Grounding of Natural Language during Human-Robot Interactions
Oliver Roesler 35

Leveraging Multimodal Behavioral Analytics for Automated Job Interview Performance Assessment and Feedback
Anumeha Agrawal, Rosa Anil George, Selvan Sunitha Ravi, Sowmya Kamath S and Anand Kumar 46

Audio-Visual Understanding of Passenger Intents for In-Cabin Conversational Agents
Eda Okur, shachi H Kumar, Saurav Sahay and Lama Nachman 55

AI Sensing for Robotics using Deep Learning based Visual and Language Modeling
yuvaram singh and Kameshwar Rao JV 60

Exploring Weaknesses of VQA Models through Attribution Driven Insights
Shaunak Halbe 64

Conference Program

A Transformer-based joint-encoding for Emotion Recognition and Sentiment Analysis
Jean-Benoit Delbrouck, Noé Tits, Mathilde Brousmiche and Stéphane Dupont

A Multi-modal Approach to Fine-grained Opinion Mining on Video Reviews
Edison Marrese-Taylor, Cristian Rodriguez, Jorge Balazs, Stephen Gould and Yutaka Matsuo

Multilogue-Net: A Context-Aware RNN for Multi-modal Emotion Detection and Sentiment Analysis in Conversation
Aman Shenoy and Ashish Sardana

Low Rank Fusion based Transformers for Multimodal Sequences
Saurav Sahay, Eda Okur, shachi H Kumar and Lama Nachman

Unsupervised Online Grounding of Natural Language during Human-Robot Interactions
Oliver Roesler

Leveraging Multimodal Behavioral Analytics for Automated Job Interview Performance Assessment and Feedback
Anumeha Agrawal, Rosa Anil George, Selvan Sunitha Ravi, Sowmya Kamath S and Anand Kumar

Audio-Visual Understanding of Passenger Intents for In-Cabin Conversational Agents
Eda Okur, shachi H Kumar, Saurav Sahay and Lama Nachman

AI Sensing for Robotics using Deep Learning based Visual and Language Modeling
yuvaram singh and Kameshwar Rao JV

Exploring Weaknesses of VQA Models through Attribution Driven Insights
Shaunak Halbe