

SHIZHE CHEN

Email: cszhe1@ruc.edu.cn

Website: <https://cshizhe.github.io>

RESEARCH INTEREST

Vision-and-Language, Video Understanding, Affective Computing, Multimodal Machine Learning

EDUCATION

Renmin University of China

Ph.D in Computer Science

Sep. 2015 - Present

Advisor: Qin Jin

Renmin University of China

B.S in Computer Science

Sep. 2011 - Jun. 2015

Advisor: Qin Jin

RESEARCH EXPERIENCE

University of Adelaide, Visiting Scholar

Image caption generation.

Jul. 2019 - Oct. 2019

Advisor: Qi Wu

Microsoft Research Asia, Research Intern

Neural storyboard generation.

Dec. 2018 - Jun. 2019

Mentors: Jianlong Fu and Ruihua Song

Carnegie Mellon University, Visiting Scholar

Video caption generation.

Oct. 2017 - Oct. 2018

Advisor: Alexander Hauptmann

RESEARCH PROJECTS

Visual Caption Generation

- Short video captioning: proposed topic-guided video captioning models published in ICMR, ACM MM and TMM; winner of MSR-VTT 2016-2017, Trecvid VTT 2016-2019.
- Dense video captioning: proposed event proposal generation and contextual-aware event caption generation system; winner of CVPR ActivityNet Dense Captioning 2018-2019.

Visual-pivoted Zero-resource Machine Translation

- Bilingual lexicon induction: proposed to employ grounded visual representation and word contextual representation learned from multilingual image caption datasets to translate words.
- Sentence translation: propose a progressive learning framework which firstly grasps word translation from image-pivots and then progresses to more challenging zero-resource sentence translation.

Multimodal Emotion Recognition

- Multimodal fusion: proposed a conditional multimodal fusion model which dynamically attends to more convincing modalities in different situations for emotion recognition.
- Emotion in dyadic dialogs: proposed to dynamically employ both self and interlocutor's contexts to improve emotion understanding in dialogs; winner of ACM Multimedia AVEC 2017-2019.
- Cross-culture emotions: proposed an adversarial framework to transfer the emotion recognition ability from one culture to another culture without sufficient training data.

TECHNICAL SKILLS

**Programming
Framework**

Python, C/C++, Javascript, MATLAB, Bash
Pytorch, Tensorflow

AWARDS

★ Ranked 1st in CVPR ActivityNet Dense Video Captioning Challenge.	2018 - 2019
★ Ranked 2nd in ICCV VATEX Video Captioning Challenge.	2019
★ Ranked 1st in NIST Trecvid Video to Text Challenge.	2017 - 2019
★ Ranked 1st in ACM Multimedia AVEC Emotion Recognition Challenge.	2018 - 2019
★ Ranked 1st in ACM Multimedia Video to Language Grand Challenge.	2016 - 2017
★ Ranked 2nd in ACM Multimedia AVEC Emotion Recognition Challenge.	2016
★ Ranked 1st in MediaEval Emotion Impact of Movies Task.	2016
★ Ranked 2nd in CCPR Multimodal Emotion Recognition Challenge.	2016
★ Second Prize in Chinese Big Data Contest P2P Task.	2015
★ Second Prize in IBM Bleumix Cognitive Computation Development Contest (developed an APP for learning English)	2015
★ First Prize in National College Student Information Security Contest.	2014
★ Second Prize in Chinese Big Data Contest Baidu IErMu Task.	2014
★ Meritorious Winner in American Mathematical Contest in Modeling.	2014
★ National Second Prize in China Undergraduate Mathematical Contest in Modeling.	2013

HONORS

★ ACM Multimedia Student Travel Grant.	2019
★ ICMR Best Paper Runner up.	2018
★ Baidu Scholarship (10 Ph.D student worldwide).	2017
★ National Scholarship for Ph.D Students.	2016
★ ACM Multimedia Student Travel Grant.	2016
★ National Scholarship for Undergraduate Students.	2013

PUBLICATION

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.