Yuxuan WANG

Mobile: 65-82252791/86-13153192997 |E-mail: yuxuan.www@gmail.com

EDUCATION

Beihang University **09/2016-07/2020**

- ***Bachelor's Degree in Electronic Information Engineering with GPA: 87.5/100 (First-Class)***
- ***Honors:*** The Second-Class Prize on China Undergraduate Physics Tournament, Beijing Division; the First-class Scholarship on Academic Competition of Beihang; the First-class Scholarship of Literature and Art Award; the Second-class Scholarship on Academic Performance

National University of Singapore (NUS) **08/2021-05/2022**

- ***Master of Science in Electrical Engineering with GPA: 4.47/5***
- ***Supervised by Prof. Mike Zheng Shou***
- ***Service: As External Reviewer of ICME 2022***

Nanyang Technological University (NTU) **08/2022-Present**

- ***Doctor of Philosophy in Computer Science and Engineering***
- ***Supervised by Prof. Hanwang Zhang***

LANGUAGE

- ***GRE: 331(Verbal: 161 / Quantitative: 170), 4.0 in Analysis Writing***
- ***TOEFL: 110(Reading: 30 / Listening: 30 / Speaking: 23 / Writing: 27)***

GRADUATE RESEARCH

Generic Event Boundary Captioning: A New Benchmark for Status Changes Understanding ***08/2021-03/2021***

Author: **Yuxuan Wang**, Difei Gao, Licheng Yu, Stan Weixian Lei, Matt Feiszli, Mike Zheng Shou

Arxiv: https://arxiv.org/abs/2204.00486

Under Reviewing of ECCV 2022

Content: In this paper, we introduce a new dataset called Kinetic-GEBC (Generic Event Boundary Captioning). The dataset consists of over 170k boundaries associated with captions describing status changes in the generic events in 12K videos. Upon this new dataset, we propose three tasks supporting the development of a more fine-grained, robust, and human-like understanding of videos through status changes. We evaluate many representative baselines in our dataset, where we also design a new TPD (Temporal-based Pairwise Difference) Modeling method for visual difference representation and achieve significant performance improvements.

AssistSR: Affordance-centric Question-driven Video Segment Retrieval ***08/2021-Present***

Author: Stan Weixian Lei, **Yuxuan Wang**, Dongxing Mao, Difei Gao, Mike Zheng Shou

Aimed at EMNLP 2022

Arxiv: https://arxiv.org/abs/2111.15050

Content: We construct a new benchmark including a dataset and a new task called Affordance-centric Question-driven Video Segment Retrieval (AQVSR), which aims at retrieving affordance-centric instructional video segments given users' questions. This dataset aims to contain 10k multimodal questions on video segments from instructional videos on diverse daily-used items. To address the task, we developed a straightforward model called Dual Multimodal Encoders (DME), which outperformed all other related methods in our task.

UNDERGRADUATE RESEARCH

Test System Building of IMS-MLD Decoding Algorithm and Reed-Muller Code ***02/2020-06/2020***

Thesis With Prof. Qin Huang, Beihang University

Content: The aim of this project is to build a test system that could reflect the SNR performance of IMS-MLD Algorithm, which is a new decoding algorithm of Reed-Muller Code.

Analog Fountain Code (AFC) ***07/2019-09/2019***

With Prof. Qin Huang, Beihang University

Content: Based on the structure of LT code, combine modulation with AFC to obtain a flexible and reliable adaptive coded modulation scheme. More than 1dB link gain and more than 30% system throughput improvement are achieved in the NG-BH scenario. This project is a business project that cooperated with the enterprise.

Blind Identification and Demodulation of Modulated Signals ***07/2018-10/2018***

With Prof. Qin Huang, Beihang University

Content: The purpose is to obtain the control signals of small UAVs around the civil aviation airport and conduct manual control to ensure the flight safety of the airport (mainly in civil aviation safety, non-military). The first step is to intercept the control signals of drones and compare and identify the 14 standard civil modulation signals one by one under the circumstance of without knowing the modulation mode to obtain the modulation mode of the signals.

PROFESSIONAL EXPERIENCE

Big Data Platform for Health and Medical, Inspur Co., Ltd ***06/2020-09/2020***

Java Software Engineer

Zhiyue, Weitong Education Co., Ltd ***10/2020-01/2021***

C# Web Software Engineer

Designing of Financial System of COMAC, Inspur Co., Ltd ***01/2021-05/2021***

Back-End Software Engineer

EXTRACURRICULAR ACTIVITIES

Student Chorus, Beihang University ***09/2016-06/2020***

Tenor

- Won the gold medal at the national college music performance
- Won the gold medal at the International Chorus Festival in Calella, Spain
- Performed in five special concerts and musical theaters
- Jointly performed with the Chorus of University of Johannesburg

SKILLS

- Computer Vision Related: PyTorch and related packages
- Other Computer Science Related: Data Structure, Algorithm, Operating System, Computer Organization and Design
- Programming Languages: Java and Web programming, SQL, Python, C# and ASP.NET MVC, C++, Verilog, Assembly Language
- Software and Tools: MATLAB, MySQL, Xilinx ISE, Vivado, Multisim, ModelSim
- Music skills: Piano, Guitar, Clarinet, Cavalry trumpet and Singing (profound experience in chorus and a cappella, also served as lead vocal in band), amateur composition and arrangement, Chorus Conducting