

JIE LEI

Email: jielei@cs.unc.edu
Phone: 919-564-8651

Updated Jan 17, 2022
Website: <https://www.cs.unc.edu/~jielei/>
Github: <https://github.com/jayleicn>





EDUCATION

- University of North Carolina at Chapel Hill** Aug 2017 - Expected May 2022
Ph.D. in Computer Science
Advisors: Tamara L. Berg and Mohit Bansal
- University of Electronic Science and Technology of China (UESTC)** Sep 2013 - Jun 2017
B.Eng in Computer Science & Technology, Yingcai Honors College
Advisor: Sinno Jialin Pan (NTU, Singapore)

EXPERIENCE

- University of North Carolina at Chapel Hill** Aug 2017 - Present
Research Assistant with Tamara L. Berg and Mohit Bansal
Chapel Hill, NC
- Vision and language understanding, esp. video and language.
 - VLEP dataset for video-and-language future event prediction. (EMNLP 2020)
 - TVR and TVC datasets for video-subtitle moment retrieval and captioning. (ECCV 2020)
 - TVQA and TVQA+ for large-scale spatio-temporal video question answering. (EMNLP 2018, ACL 2020)
- Facebook AI** Aug 2021 - Dec 2021
Part-time Student Researcher with Licheng Yu, Xinlei Chen and Ning Zhang
Menlo Park, CA
- Vision and language, especially image retrieval.
- Facebook AI** May 2021 - Aug 2021
Research Scientist Intern with Licheng Yu, Xinlei Chen and Ning Zhang
Menlo Park, CA
- Vision and language, especially image retrieval.
- Microsoft Cloud & AI** May 2020 - Aug 2020
Researcher Intern with Linjie Li, Luowei Zhou, Zhe Gan and Jingjing Liu
Bellevue, WA
- End-to-end vision and language pretraining. (CVPR 2021)
- Tencent AI Lab** May 2019 - Aug 2019
Research Intern with Liwei Wang, Yelong Shen and Dong Yu
Bellevue, WA
- Recurrent transformer for coherent video paragraph captioning. (ACL 2020)
- Nanyang Technological University** Oct 2016 - Apr 2017
Research Intern with Sinno Jialin Pan
Singapore
- Transfer learning in deep hash models for fast image retrieval.
- University of Manitoba** Jul 2016 - Oct 2016
Research Intern with Yang Wang
Winnipeg, Canada
- Weakly supervised image classification using hierarchical image labels. (CRV 2017)

PUBLICATIONS

- [14] Hao Tan*, **Jie Lei***, Thomas Wolf, Mohit Bansal, “VIMPAC: Video Pre-Training via Masked Token Prediction and Contrastive Learning”, arXiv 2021.  Paper,  Code
- [13] **Jie Lei**, Tamara L. Berg, Mohit Bansal, “QVHighlights: Detecting Moments and Highlights in Videos via Natural Language Queries”, NeurIPS 2021.  Paper,  Code

- [12] Linjie Li*, **Jie Lei***, Zhe Gan, Licheng Yu, Yen-Chun Chen, Rohit Pillai, Yu Cheng, Luowei Zhou, Xin Eric Wang, William Yang Wang, Tamara L. Berg, Mohit Bansal, Jingjing Liu, Lijuan Wang, Zicheng Liu, “VALUE: A Multi-Task Benchmark for Video-and-Language Understanding Evaluation”, NeurIPS 2021 Datasets and Benchmarks Track. 📄 Paper, 🏆 Leaderboard
- [11] Linjie Li, **Jie Lei**, Zhe Gan, Jingjing Liu, “Adversarial VQA: A New Benchmark for Evaluating the Robustness of VQA Models”, ICCV 2021 **Oral**. 📄 Paper, 🗃️ Dataset
- [10] **Jie Lei**, Tamara L. Berg, Mohit Bansal, “mTVR: Multilingual Moment Retrieval in Videos”, ACL 2021. 📄 Paper, 🔄 Code
- [9] Jaemin Cho, **Jie Lei**, Hao Tan, Mohit Bansal, “Unifying Vision-and-Language Tasks via Text Generation”, ICML 2021. 📄 Paper, 🔄 Code
- [8] Zineng Tang*, **Jie Lei***, Mohit Bansal, “Improved Pre-Training from Noisy Instructional Videos via Dense Captions and Entropy Minimization”, NAACL 2021. 📄 Paper, 🔄 Code
- [7] **Jie Lei***, Linjie Li*, Luowei Zhou, Zhe Gan, Tamara L. Berg, Mohit Bansal, Jingjing Liu, “Less is More: Clip-BERT for Video-and-Language Learning via Sparse Sampling”, CVPR 2021 **Best Student Paper Honorable Mention, Oral**. 📄 Paper, 🔄 Code
- [6] **Jie Lei**, Licheng Yu, Tamara L. Berg, Mohit Bansal, “What is More Likely to Happen Next? Video and Language Future Event Prediction”, EMNLP 2020. 📄 Paper, 🗃️ Dataset
- [5] **Jie Lei**, Licheng Yu, Tamara L. Berg, Mohit Bansal, “TVR: A Large-Scale Dataset for Video-Subtitle Moment Retrieval”, ECCV 2020. 📄 Paper, 🗃️ Dataset, 🔄 Retrieval Code, 🔄 Captioning Code
- [4] **Jie Lei**, Liwei Wang, Yelong Shen, Dong Yu, Tamara L. Berg, Mohit Bansal, “MART: Memory-Augmented Recurrent Transformer for Coherent Video Paragraph Captioning”, ACL 2020. 📄 Paper, 🔄 Code
- [3] **Jie Lei**, Licheng Yu, Tamara L. Berg, Mohit Bansal, “TVQA+: Spatio-Temporal Grounding for Video Question Answering”, ACL 2020. 📄 Paper, 🗃️ Dataset, 🔄 Code
- [2] **Jie Lei**, Licheng Yu, Mohit Bansal, Tamara L. Berg, “TVQA: Localized, Compositional Video Question Answering”, EMNLP 2018 **Oral**. 📄 Paper, 🗃️ Dataset, 🔄 Code
- [1] **Jie Lei**, Zhenyu Guo, Yang Wang. “Weakly Supervised Image Classification with Coarse and Fine Labels.” CRV 2017. 📄 Paper, 🔄 Code

PROJECTS

- [1] **Jie Lei**, “AnimeGAN: Create Anime Face using Generative Adversarial Networks”, GitHub 2017.
- Implementation of GANs for anime face generation. 🔄 Data & Code (**1K+** stars)
 - Automatic pipeline for large-scale anime face dataset construction from the web.

AWARDS

Best Student Paper Honorable Mention , CVPR 2021, Virtual	2021
Adobe Research Fellowship , Adobe, USA	2021
Outstanding Undergraduate Thesis , UESTC, China	2017
Globalink Research Intern Award (200 undergraduates across China), Mitacs, Canada.	2016

ACADEMIC SERVICE

Journal and Conference Reviewer: TPAMI, IJCV, CVPR 2019-2022, ICCV 2019, ECCV 2020, ACM MM 2020 Interactive Arts, AAAI 2021, ICCV 2021, ACL 2021-2022, NeurIPS 2021, ICLR 2021

Workshop Program Committe: NeurIPS HAMLETS 2020