# Tian Yun

tttyuntian.github.io | tian_yun@brown.edu | (530) 551-9835

## EDUCATION

**Brown University** Providence, RI
*Doctor of Philosophy in Computer Science* *September 2022 - Present*

**Brown University** Providence, RI
*Master of Science in Computer Science* *August 2020 - May 2022*

**Wake Forest University** Winston-Salem, NC
*Bachelor of Science in Mathematical Statistics* *August 2016 - May 2020*
*Honors Bachelor of Science in Computer Science*

## PUBLICATIONS

[1] *mOthello: When Do Cross-Lingual Representation Alignment and Cross-Lingual Transfer Emerge in Multilingual*
Tianze Hua*, **Tian Yun***, Ellie Pavlick
*NAACL 2024*

[2] *Emergence of Grounded Representations in Embodied Sequence Modeling*
**Tian Yun***, Zilai Zeng*, Kunal Handa, Ashish Thapliyal, Bo Pang, Ellie Pavlick, Chen Sun
*EMNLP 2023*

[3] *Improved Inference of Human Intent by Combining Plan Recognition and Language Feedback*
Ifrah Idrees, **Tian Yun**, Naveen Sharma, Nakul Gopalan, Stefanie Tellex, George Konidaris
*IROS 2023*

[4] *Do Vision-Language Pretrained Models Learn Primitive Concepts?*
**Tian Yun**, Usha Bhalla, Ellie Pavlick, Chen Sun
*TMLR*

[5] *Does Vision-and-Language Pretraining Improve Lexical Grounding?*
**Tian Yun**, Chen Sun, Ellie Pavlick
*Findings of EMNLP, 2021*

[6] *Mining Biomedical Texts for Pediatric Information*
**Tian Yun**, Deepti Garg, Natalia Khuri
*14th International Joint Conference on Biomedical Engineering Systems and Technologies, 2021*

## PREPRINTS

[7] *BLOOM: A 176B-Parameter Open-Access Multilingual Language Model*
Teven Scao et al.
*arXiv preprint arXiv:2211.05100 (2022)*

## RESEARCH EXPERIENCE

**LUNAR Lab & PALM Lab, Brown University** Providence, RI
*Co-advised by Ellie Pavlick and Chen Sun* *August 2020 – Present*

- Investigated whether vision-and-language (VL) pretraining yields better linguistic representations.
- Researched whether pretrained vision-and-language (VL) models learn composable primitive concepts.
- Studied how grounded world representations emerge in embodied sequence modeling.

**Humans to Robots Lab, Brown University** Providence, RI
*Advised by Stefanie Tellex* *March 2021 – December 2023*

- Proposed a situation-aware coaching dialogue system that guides users to accomplish daily tasks by reasoning over uncertainties in the environment and asking clarification questions to the users.

**Google** Mountain View, CA
*Student Researcher, Advised by Bo Pang and Ashish Thapliyal* *June 2022 – August 2022*

- Studied how well vision-and-language models can leverage learned visual commonsense knowledge to adapt to new visual commonsense knowledge in synthetic environment.

**BigScience Workshop, HuggingFace** Providence, RI
*Advised by Ellie Pavlick* *May 2021 – Aug 2022*

- Co-led the engineering team in Evaluation working group to build an evaluation pipeline with a large set of prompted datasets for large language models.

**DataMine Research Group, Wake Forest University** Winston-Salem, NC
*Advised by Natalia Khuri* *August 2019 – August 2020*

- Built a pipeline with machine learning algorithms to extract information about safety and efficacy of drugs in pediatric populations from approved drug product labels.

## TEACHING ASSISTANTSHIPS

**CSCI-1470/2470: Deep Learning, Brown University** *September 2021 – December 2021*
*Teaching Assistant* Instructor: Chen Sun

**CSCI-1460: Computational Linguistics, Brown University** *January 2021 – April 2021*
*Teaching Assistant* Instructor: Eugene Charniak

**Computer Science Peer Tutor, Wake Forest University** *September 2017 – May 2020*

## SERVICE (*outstanding reviewer)

**Conference Reviewer**

- Computer Vision and Pattern Recognition (CVPR) *2022, 2023, 2024*
- European Conference on Computer Vision (ECCV) *2022, 2024*
- International Conference on Computer Vision (ICCV) *2023**