

Li Sun

🔗 lisun-ai.github.io 🌐 github.com/lisun-ai ✉ lisun@bu.edu ☎ (412)-209-5278

RESEARCH INTERESTS

Machine learning: Un-/Self-supervised learning, Adversarial learning, Generative models
Natural language processing: Large language model, Model robustness

EDUCATION

Boston University Jan 2023 - May 2024 (*Expected*)
PhD student in Computer Engineering Boston, MA
Advisor: Prof. Kayhan Batmanghelich

University of Pittsburgh Aug 2019 - Dec 2022
PhD student in Intelligent Systems (Transferred to BU) Pittsburgh, PA
M.Sc. in Intelligent Systems (Artificial Intelligence)
GPA: 3.93/4.00, advisor: Prof. Kayhan Batmanghelich

SELECTED PUBLICATIONS

Sun, L., Luisier, F., Batmanghelich, K., Florencio, D., & Zhang, C. (2023).
From Characters to Words: Hierarchical Pre-trained Language Model for
Open-vocabulary Language Understanding. In *Proceedings of the Annual Meeting
of the Association for Computational Linguistics (ACL)*. **Best Papers Honorable Mention**

Robinson, J., **Sun, L.**, Yu, K., Batmanghelich, K., Jegelka, S., & Sra, S. (2021).
Can contrastive learning avoid shortcut solutions? In *Advances in Neural Information
Processing Systems (NeurIPS)*.

Sun, L., Yu, K., & Batmanghelich, K. (2021). Context Matters: Graph-based
Self-supervised Representation Learning for Medical Images. In *Proceedings of the
AAAI Conference on Artificial Intelligence (AAAI)*.

Saeedi, A., Utsumi, Y., **Sun, L.**, Batmanghelich, K. & Lehman L. (2022). Knowledge
Distillation via Constrained Variational Inference. In *Proceedings of the AAAI
Conference on Artificial Intelligence (AAAI)*.

Sun, L., Chen, J., Xu, Y., Gong, M., Yu, K., & Batmanghelich, K. (2022). Hierarchical
Amortized GAN for 3D High Resolution Medical Image Synthesis. In *IEEE journal of
biomedical and health informatics*

WORK EXPERIENCE

Meta May 2023 - Sept 2023
Research Scientist Intern, manager: Dr. Philip Bontrager New York, NY
- Worked on leveraging large language models for improved visual representation learning.
- Developed a token-level visualization method for explanation of large multi-modal models.
- Studied the robustness and generalizability of visual representation from large
multi-modal models.

Microsoft June 2022 - Aug 2022
Research Intern, manager: Dr. Dinei Florencio Bellevue, WA
- Designed and implemented a hierarchical language model for open-vocabulary
language understanding
- Experiments demonstrated that our model outperforms baselines, and it is more
robust to textual corruptions and domain shifts
- Manuscript accepted by ACL

Microsoft

Research Intern, manager: Dr. Eric I-Chao Chang

Mar 2019 - Aug 2019

Beijing, China

- Developed fine-grained model that makes use of anatomical structure for chest radiograph interpretation
- Achieved improved results on thoracic diseases that are subtle and require close observation
- Developed reinforcement learning model for interpretable life support device detection

TECHNICAL SKILLS

Languages: Python, R, Shell, Java, C++

Softwares: PyTorch, Tensorflow

RESEARCH EXPERIENCE

Avoiding Shortcut Solutions in Contrastive Learning

Nov 2020 - May 2021

Mentor: Prof. Suvrit Sra, MIT & Prof. Kayhan Batmanghelich, University of Pittsburgh

- Analyzed feature suppression in contrastive learning, and explained why feature suppression can occur when optimizing the InfoNCE loss
- Proposed implicit feature modification, a simple and efficient method that reduces the tendency to use feature suppressing shortcut solutions and improves generalization
- Highlighted by World Economic Forum and MIT News

Hierarchical Language Model for Open-vocabulary Language Understanding

June 2022 - Dec 2022

Mentor: Florian Luisier & Dinei Florencio, Microsoft

- Introduced a novel pre-trained language model with hierarchical two-level architecture: one at the word level and another at the sequence level
- Demonstrated that our hierarchical model attains better performance on noisy text and cross-domain tasks

Context-aware Self-supervised Learning for Medical Images

May 2020 - Sept 2020

Mentor: Prof. Kayhan Batmanghelich, University of Pittsburgh

- Proposed a self-supervised representation learning method for volumetric medical images that accounts for anatomical context, which is from both local anatomical profiles and graph-based relationship
- Proposed method that provides task-specific explanation for the predicted outcome
- Short version accepted by Medical Imaging meets NeurIPS workshop (Oral), long version paper accepted by AAAI

Hierarchical Amortized GAN for 3D High Resolution Medical Image Synthesis

Jan 2020 - Nov 2020

Mentor: Prof. Kayhan Batmanghelich, University of Pittsburgh

- Proposed a novel end-to-end GAN architecture that can generate high-resolution volumetric images while being memory efficient
- Discovered that moving along certain directions in latent space results in explainable anatomical variations in generated images
- Paper accepted by IEEE journal of biomedical and health informatics

Brain Tumor Segmentation and Survival Prediction with Deep Learning

May 2018 - Aug 2018

Mentor: Prof. Lin Luo, Peking University

- Developed an ensemble model of 3D CNNs to segment brain tumor from multimodal MRI scans, then extracted radiomic features from segmented tumor combined with clinical features to predict patients' overall survival
- Ranked 2nd place and 5th place out of 60+ teams in 2018 MICCAI BraTS challenge on survival prediction task and segmentation task respectively, received prize from Intel AI
- Paper accepted by MICCAI BrainLes 2018 workshop (Spotlight)