Haotian Xue

Senior Student @ CS. SJTU

Email : xavihart@sjtu.edu.cn
Mobile : +86 19921872686
Homepage : https://htxue.info

Education

- **Shanghai Jiao Tong University** — Shanghai, China
  B.E. in Computer Science — *Sept. 2018 – June 2022 (Expected)*
  Major GPA: 90.68/100 (maths related: 91.4/100)

- **Shanghai Jiao Tong University Zhiyuan Honor Class** — Shanghai, China
  Zhiyuan Honor class of Engineering(top 5%) — *Sept. 2018 – June 2022 (Expected)*

Research Interest

- **Robust and Generalized Robot Learning**

  I am interested in utilizing deep learning methods in vision and language to help robots better feel and work in the real world in a more generalized way. Specifically, I hope to make the robot able to learn physics or policies from large amount of real-world data from different modalities lack of labeling like an infant.

- **Explainable Artificial Intelligence**

  I am interested in explaining some deep learning phenomenons like overfitting and information redundancy with mathematical foundations like game theory. From both model-centric and data-centric angles, I hope that we can get a better grasp of the DNN model, to better evaluate and even deug them.

Research Experience

- **Massachusetts Institute of Technology, CSAIL** — remote intern
  Intuitive Physics, *Research Intern, advised by Prof. Josh Tenenbaum* — *Jul 2021 – Now*
  - **[Visual Dynamics]**
    Focus on improving generalization on visual intuitive physics. The key point of this work is to learn dynamics from video inputs in a compositional and interpretable way. Also, we try to make the whole pipeline adapt to local-varied scenarios to make it generalized. Also it will help with robot manipulation, especially for scenarios with fluid.
    [One **first-author** paper in preparation]

- **John Hopcroft Center for Computer Science** — Shanghai, China
  Natural Language Processing, *Research Intern, tutored by Prof .Zhouhan Lin* — *May 2021 – Now*
  - **[Syntactically Enhanced Transformer]**
    Using the constituency parsing result of source texts to help with 1. bring syntactic distance into transformer to get an interpretable and dynamic attention 2. boost performance of transformer used for NLP tasks like machine translation and text summary.
    [One **first-author** paper in preparation]
  - **[Distance-Transformer]**
    Bring syntactic distance into transformer which will help with unsupervised learning of syntactic information and will in turn help with learning interpretable self-attention weight and boost performance.
    [One **first-author** paper in preparation]

- **John Hopcroft Center for Computer Science** Shanghai, China
  Explainable AI / Game Theory, *Research Intern, tutored by Prof .Quanshi Zhang* *Jun 2019 – May 2021*
  - **[Evaluation for Explanation Methods]**
    Work on proposing a set of criteria to evaluate the objectiveness of explanation methods (like LRP, CAM, grad-CAM, Gradient-Input) of neural networks based on Shapley value in game theory. We prove that the method we proposed is effective and stable.
    [One **second-author** paper in submission]
  - **[Generalization Problem and Shapley Interaction]**
    Work on explaining for generalization problem in deep learning based on multi-player interaction in game theory. We build a synthetic dataset based on CUB to calculate the shapley interaction between different part of the bird body like head, leg and wings.
  - **[Explaining and Boosting Deep Prototypical Concepts]**
    Work on proposing a hypothetical explanation for the cognitive difficulty of an image, using the Harsanyi dividend to explain how a pre-trained DNN represents prototypical concepts and non-prototypical concepts, and further design a set of image revision operations to decrease the cognitive difficulty of an image.
    [One **second-author** paper in submission]

INTERNSHIP

- **Microsoft Research Asia, Natural Language Computing Group** Beijing, China
  Document Intelligence, *Research Intern, tutored by Dr .Lei Cui* *Sept 2021 – Dec 2021*
  - **[Document Vision Transformer]**
    Work on large-scale pre-trained ViT-based model for document AI. We try to bring pre-trained models into document AI to help with learning a better representation for document images, which will boost performance in down-streaming document understanding tasks like OCR detection and layout analysis.

Paper in Submissions/Preparation

Hao Zhang, **Haotian Xue**, Jiayi Chen and Quanshi Zhang, *Towards a Unified Evaluation of Explanation Methods without Ground Truth* (In submission to WWW2022)

Xu Cheng*, Xin Wang*, **Haotian Xue**, Zhengyang Liang and Quanshi Zhang, *A Hypothesis for the Cognitive Difficulty of Images* (In submission to CVPR2022)

Haotian Xue, Yunzhu Li, Hsiao-Yu Tung, Antonio Torralba, Daniel LK Yamins, Joshua B. Tenenbaum, *A Compositional 3D Representations for Improving Generalization in Visual Intuitive Physics and Manipulation* (In prepration)

Haotian Xue, Shengyuan Hou, Zhouhan Lin, *Neural Machine Translation enhanced by supervised syntactic distance* (In prepration)

Awards

- **Zhiyuan College Honors Scholarship 2018,2019,2020,2020 CN ¥20000, Top5%)**:
- **Singapore Technologies Engineering Scholarship (CN ¥5000, Top 20%)**:
- **Zhiyuan College Excellent Scholarship (Class C, CN ¥2000)**: