

# Jinfa Huang

✉ jinfahuang@stu.pku.edu.cn

🌐 <https://infaaa.github.io/>

☎ +86 18319900570

## Education

### 📖 Peking University, China

[Sep. 2020 – Jul. 2023 (Expected)]

*Master in Computer Science*

- Courses: Digital Media Software (A+), Medical Image Processing (A+), Artificial Intelligence (A+), Computer Vision (A+).
- GPA: 3.72/4.00, advised by Prof. [Jie Chen](#).

### 📖 University of Electronic Science and Technology of China, China

[Sep. 2016 – Jun. 2020]

*Bachelor of Software Engineering*

- Honor&Award: Outstanding Graduate, Best debater.
- GPA: 89.44/100; GPA Ranking: [5/171 \(3.0%\)](#)

## Research Interests

- 📖 I aim at building multimodal interactive AI systems that can not only ground and reason over the external world signals, e.g., vision, audio and knowledge, to understand human language, but also assist humans in decision-making and efficiently solving social concerns, e.g., robot. As steps towards this goal, my research interests include but are not limited to **Visual Captioning**, **Video-Text Retrieval** and **Visual Question Answering**.

## Research Experiences

### 📖 Noah's Ark Lab, Huawei, Shenzhen, China.

[Apr. 2022 – Present]

*Research Intern*, working on visual prompt tuning.

- Advisors: Prof. [Jianzhuang Liu](#), and Dr. [Yi Zhu](#).
- Brief introduction: We conducted research on visual prompt tuning and prototypical contrastive learning for video captioning. i) Research on how to contrast prompt tuning between hard and soft prompts to better adapt for downstream visual tasks. ii) Research prototypical contrastive learning for the generative task and use the Expectation-Maximization Algorithm to demonstrate the method's validity.

### 📖 Medical Imaging Group, Pengcheng Lab, Shenzhen, China.

[Sep. 2020 – March. 2022]

*Research Intern*, working on AI in Healthcare.

- Advisor: Prof. [Jie Chen](#), and Dr. [Guoli Song](#).
- Brief introduction: We conducted research on visual captioning and medical report generation, mainly studying visual captioning under complex visual data. (i) Research on how to use multi-modal pre-trained models (e.g., CLIP) for facilitating visual captioning. (ii) Research on X-Ray images in COVID-19, reduce the pressure on doctors for diagnosis by automatically generating medical reports, and propose a memory-driven transformer for the abnormal regions.

### 📖 Multimedia Computing Team, KDDI Research, Saitama-Shi, Japan.

[Nov. 2019 – Feb. 2020]

*Research Intern*, working on group cohesion recognition and multi-modal emotion classification.

- Advisor: Dr. [Jianming Wu](#), and Dr. [Yanan Wang](#).
- Brief introduction: We conducted research on group cohesion recognition and emotion classification, mainly studying the visual and language fusion and representation, helping improve related emotion analysis tasks and human-computer interaction landing.
- Assist in the design of semantic embedding-related modules for multi-modal data fusion of images and texts to condense the knowledge in the text to improve the effect of group cohesion recognition. The classification result by MSE in the test set has been reduced by about 0.06.
- Participate in the competition of task 8 in SemEval 2020. By using BERT, ELMo, BiGRU, Resnet, and fusion of five models to integrate the hybrid model. Tasks: sentiment classification, humor classification, and fine-grained semantic classification, where we achieve the performance of all in the top 2 ranking.

### 📖 X-Data Research Group, Tencent IEG, Shenzhen, China.

[Jan. 2019 – Jul. 2019]

*Engineering Intern*, working on text classification and information extraction in Game AI.

- Mentor: Dr. [Yang Chao](#).
- Brief introduction: We are mainly responsible for game public sentiment analysis, game context classification and filtering, and related information extraction, using BERT for sentiment analysis. We aim to serve players, enhance the gambling experience and reduce manual pressure.

## Honors & Awards

- 📌 Excellent Student Scholarship. (Top 10%) [2017 2018 and 2019]
- 📌 National Inspirational Scholarship, UESTC, China. [2018]
- 📌 Bronze, ACM-CCPC, China Collegiate Programming Contest, Jilin. [2018]
- 📌 Honorable Mention, Mathematical Contest in Modeling (MCM). [2018]
- 📌 Selected entrant for 2019 Google Machine Learning Winter Camp (100 participants nationwide). [2019]
- 📌 Selected entrant for 2020 Deep Learning Summer Camp (200 participants nationwide). [2020]
- 📌 Ranking: 25/3219 (0.7%), Leetcode Spring LCCUP. [2021]

## Publications

- [6] 📌 *Knowledge Bottleneck Distillation for Grounded Image Captioning*
  - **Jinfa Huang**, Hao Li, Guoli Song, Jie Chen.
  - In Proceedings of the Empirical Methods in Natural Language Processing ([EMNLP](#)), 2022. (under review)
- [5] 📌 *Expectation-Maximization Contrastive Learning for Compact Video-and-Language Representations*
  - Peng Jin\*, **Jinfa Huang\***, Fenglin Liu, Shen Ge, Guoli Song, Xian Wu, Jie Chen.
  - In Proceedings of the Neural Information Processing Systems ([NeurIPS](#)), 2022. (under review)
- [4] 📌 *Aligning Disentangled Representation via Adaptive Centers for Text-Video Retrieval*
  - Jin Peng\*, **Jinfa Huang\***, Hao Li, Li Yuan, Guoli Song, Jie Chen.
  - In Proceedings of the Neural Information Processing Systems ([NeurIPS](#)), 2022. (under review)
- [3] 📌 *Beyond 2D: 3D Spatial Reasoning for TextVQA*
  - Hao Li\*, **Jinfa Huang\***, Peng Jin, Guoli Song, Qi Wu, Jie Chen.
  - In Proceedings of the ACM International Conference on Multimedia ([ACMMM](#)), 2022. (under review)
- [2] 📌 *LDNN: Linguistic Knowledge Injectable Deep Neural Network for Group Cohesiveness Understanding*
  - Yanan Wang, Jianming Wu, **Jinfa Huang**, Gen Hattori, Jie Chen, Satoshi Kurihara.
  - In Proceedings of the ACM International Conference on Multimodal Interaction ([ICMI](#)), 2020.
- [1] 📌 *Guoym at SemEval-2020 Task 8: Ensemble-based Classification of Visuo-Lingual Metaphor in Memes*
  - Yingmei Guo, **Jinfa Huang**, Yanlong Dong, Mingxing Xu.
  - Proceedings of the Fourteenth Workshop on Semantic Evaluation ([SemEval](#)), 2020.

## Skills

- Languages** 📌 Strong reading, writing and speaking competencies for English, Mandarin Chinese, Japanese.
- Coding** 📌 Python, C++, Latex, Shell, Matlab. Familiar with tensorflow and pytorch programming.
- Misc.** 📌 Academic Research, Algorithm Competition (ACM), Chinese Paladin (Game), Anime.

## Services

- 📌 **Reviewer:** IJCAI 2022, MICCAI 2022, and CVPR 2022.
- 📌 **Teaching Assistant:** Artificial Intelligence, and Medical Image Processing.