Xiaohan Zou

☎ (+1) 857-356-5018 | ✉ zxh@bu.edu | ⌨ Renovamen (3.5k stars) | 🔗 xiaohan-zou | 🌐 zxh.io

Education

Boston University   Boston, MA
M.S. in Computer Science   09/2021 – 01/2023
Tongji University   Shanghai, China
B.Eng. in Software Engineering   09/2016 – 07/2020

Publications and Submitted Manuscripts

○ **Xiaohan Zou**, Changqiao Wu, Lele Cheng, and Zhongyuan Wang. "TokenFlow: Rethinking Fine-grained Cross-modal Alignment in Vision-Language Retrieval", **submitted to** ACM International Conference on Multimedia (**ACM MM**), 2022.

○ **Xiaohan Zou**, and Tong Lin. "Efficient Meta-Learning for Continual Learning with Taylor Expansion Approximation", International Joint Conference on Neural Networks (**IJCNN**), 2022.

○ **Xiaohan Zou**, Cheng Lin, Yinjia Zhang, and Qinpei Zhao. "To be an Artist: Automatic Generation on Food Image Aesthetic Captioning", International Conference on Tools with Artificial Intelligence (**ICTAI**), 2020. (Acceptance Rate: 25%, **Oral Presentation**) [Paper] [Code] [Slide]

○ **Xiaohan Zou**. "A Survey on Application of Knowledge Graph", International Conference on Control Engineering and Artificial Intelligence (**CCEAI**), 2020. [Paper]

Professional Experience

Machine Learning Engineer Intern, Kuaishou Technology, Beijing   07/2021 – 04/2022

○ Devised a new model-agnostic formulation for fine-grained cross-modal semantic alignment and subsumed the recent popular works into the proposed scheme

○ Proposed a video-text retrieval method which is competitive when compared with the SoTA approaches with heavy model design by only altering the similarity function, **submitted to ACM Multimedia 2022**

○ Developed a PyTorch library for video-text retrieval which is benefiting the group members' research work

Software Engineer Intern, China Electronics Technology Group Corporation   10/2020 – 06/2021

○ Involved in building a security visualization system for Sanxingdui using Cesium and Vue, responsible for displaying 3D models and visualizing sensor data

○ Labeled ancient handwritten Chinese characters and generated OCR training files automatically using Python and Shell

○ Developed a demo for a real time person pose estimation model

Game Engineer Intern, Banana Interactive, Shanghai, China   10/2019 – 05/2020

○ Participated in the development and maintenance of 3 H5 games using JavaScript and Affinity Designer

○ Completed the first release version of a H5 game independently

○ Improved the readablity of code through code refactoring and developed a skin system, shop system and item system for the other two games

○ Ported a game packaging and deployment tool from Windows to Linux and macOS

Research Experience

Personalized Product Description Generation, Deecamp 2021   06/2021 – 08/2021

○ Incorporated product attributes, personalized information and external knowledge to T5 pre-trained model using transformer and bidirectional attention to generate interesting and informative product descriptions

○ Built the frontend of a fancy cross-platform website for interacting with our model using Vue and uni-app

○ Won the **champion of the language track** in Deecamp 2021

Meta-Learning for Continual Learning, Peking University, Beijing   08/2020 – 01/2022
Advisor: Prof. Tong Lin

- Designed an efficient method for parameter importance estimation via Taylor expansion
- Proposed a fast meta-learning algorithm for continual learning problems, which expresses the gradient of meta-update in closed-form instead of using Hessian information, **accepted by IJCNN 2022**
- Outperformed SoTA methods while optimizing much more efficient in experiments on popular benchmarks

Food Image Aesthetic Captioning, Tongji University, Shanghai [Project Page]   03/2020 – 06/2020
Advisor: Prof. Qinpei Zhao

- Proposed a novel framework consisting of a single-attribute captioning module and an unsupervised text summarization module for generating aesthetic captions for food images, **published in ICTAI 2020**
- Designed a data filtering strategy inspired by TF-IDF method for building a dataset for this new task
- Introduced two new evaluation criteria to assess the novelty and diversity of the generated captions
- Outperformed baselines and existed methods substantially in terms of diversity, novelty and coherence

Fault Diagnosis for Microservice Architectures, Tongji University, Shanghai   09/2018 – 01/2019

- Represented the microservices of a cloud platform and the causal relationships between them by a Bayesian network against the observed performance metrics dynamically using PC algorithm
- Identified the culprit microservices when an anomaly occured using random walk
- Outperformed traditional approaches with 6.56% accuracy improvement, without knowing the calling graph

Chatbot   10/2018 - 11/2018
Advisor: Dr. Fan Zhang

- Built a chatbot which can identify and analyze the intentions of users and provide stock and weather information using RasaNLU
- Achieved intention classification and named entity recognition based on SVM using spaCy and scikit-learn
- Allowed multi-turn dialogue and handling pending actions using finite state automaton
- Integrated the chatbot to WeChat and QQ for a better user interface

Semi-Supervised Machine Translation, Peking University, Beijing   07/2018 – 08/2018
Advisor: Prof. Tong Lin

- Proposed a dual learning framework based on shared hidden space to utilize the structure duality to boost the learning of two dual tasks and better regularize the model
- Designed two denoising auto-encoders consisting of encoders and decoders of two traditional Seq2Seq neural machine translators to make use of unpaired data
- Outperformed strong baselines by 1.0 - 2.9 BLEU on IWSLT'15 (English-Vietnamese) and WMT'14 (English-German), the improvement is more obvious when labeled data is little

Open-Source Projects

Flint: A toy deep learning framework implemented in Numpy from scratch [Github]   01/2021 – Present

- Implement an autograd engine, Linear, Convolution, Pooling, Flatten, RNN, Dropout and BatchNorm layers, 6 optimizers, 4 loss functions, 3 activation functions, 5 initializers and a data loader in pure Numpy
- Wrote complete documentation and comprehensive unit tests

Speech Emotion Recognition [Github]   04/2019 – 06/2019

- Implemented several models and features extracting methods for speech emotion recognition
- Achieved 7.2 - 12.2 accuracy improvement over baseline on four benchmark datasets: CASIA (Chinese), EMODB (German), SAVEE (English), and RAVDESS (English), got over 300 Github stars

Awards and Honors

Champion of Language Track, Deecamp   2021
Bronze, China Collegiate Programming Contest (CCPC)   2018

Honorable Mention, ACM International Collegiate Programming Contest (ICPC) Asia Regional   2018
Second Prize, Programming Contest of Tongji University   2017, 2018
Second Prize, China Mathematical Contest in Modeling (CUMCM)   2017, 2018
Second Prize, Programming Contest of East China Normal University   2017

Leadership and Activities

Vice Chief Technology Officer & Chief Experience Officer, Tongji Microsoft Student Club   2017 - 2019
- Participated in seminars and meetings
- Gave lectures about data structure and algorithms as well as their applications in machine learning

Vice President, Baseball and Softball Association of Tongji University (Jiading campus)   2018 - 2019
Member of Softball Team of Tongji University   2016 - 2019

Core Courses

- **Boston University:** Machine Learning, Image and Video Computing, Cryptography, Computational Tools for Data Science
- **Tongji University**: Probability and Mathematical Statistics, Discrete Mathematics, Calculus, Linear Algebra, Object-Oriented Programming, Data Structure, Algorithm, Operating Systems, Principles of Compilers, Database, SOA and Web Services, Distributed System, Web System and Technology, Design Patterns

Skills

Programming Languages: Python, JavaScript/TypeScript, HTML/CSS, Java, C/C++, MATLAB
Tools and Frameworks: Git, PyTorch, Keras, scikit-learn, Linux, Vue, React, Django, LaTeX
Languages: Chinese (native), English (proficient, TOEFL: 106, GRE: 322)