

HAOJIAN HUANG

TEL.: +86-135 0246 0463 · E-MAIL: haojianhuang927@gmail.com

Github: [JethroJames](#) | Personal Links: [wonghougin.me](#)

EDUCATION

Department of Computer Science, University of Hong Kong
MSc in Computer Science

Hong Kong, China
09/2024-Present

College of Computer Science and Technology, Harbin Engineering University
Bachelor of Engineering in Information Security

Harbin, China
09/2020-07/2024

- GPA: 86.94/100 (Ranking: 5/25)

PROFESSIONAL EXPERIENCES

TeleAI -- Intern (Supervisor: Associate Professor Chen Mulin)

02/2024-Present

- Led the development of innovative technologies in Zero-Shot Image Retrieval, Multimodal NER (REMEDY), and Video Temporal Localization (SRAM), significantly enhancing performance and interpretability, with research findings published in top conferences like ACM MM 2024 and NeurIPS 2024.

Hangzhou Innovation Institute of Beihang University -- Intern

10/2023-02/2024

- Conducted multimodal traffic prediction research on spatio-temporal graphs, compiled insightful reports, and implemented algorithms into functional code, providing a robust core codebase for project success and real-world applications.
- Contributed to the team's project by leading the development of advanced traffic prediction capabilities, facilitating multimodal data integration, and supporting efforts to patent the innovative technology for multimodal traffic prediction and management.

Big Data Development Center, Ministry of Agriculture and Rural Affairs of the People's Republic of China -- Researcher at the Comprehensive Data Analysis Department

03/2023-06/2023

- Deployed literature crawler systems and developed temporal analysis algorithms for comprehensive data analysis.
- Authored and submitted paper on "Ultra-short-term Multi-step Wind Speed Prediction for Wind Farms" to top-tier *EAAI* journal.
- Contributed to the development of reports and publications based on data analysis findings.

Key Laboratory of Multi-disciplinary Collaborative Cognitive AI Technology and Application in Heilongjiang Province -- Intern

03/2022-03/2023

- Organized team members to conduct field research and literature reviews, while directing software development and thesis writing as the leader of the National Undergraduate Innovation and Training Program.
- The primary research topics encompassed: "Auxiliary System for Real-time People Flow Spatio-temporal Information and Decision Optimization in Nucleic Acid Testing Based on Deep Learning" and "Short-term Wind Speed Forecasting Model based on an Attention-gated Recurrent Neural Network and Error Correction Strategy." The latter was submitted to Q1 Top *Energy*.

Huazhong University of Science and Technology - RA (Supervisor: Professor Yang Ruotian)

07/2021-05/2023

- Researched the Trustworthy AI in multi-view clustering, encompassing both theoretical exploration and practical application.
- Designed and implemented the model's code framework from scratch, ensuring its efficiency, maintainability, and scalability.
- Composed a high-quality academic paper based on research findings, clearly articulating research motivation, methodology, experimental design, and results analysis to demonstrate the model's innovation and practicality.

RESEARCH PROJECTS

CREST: Cross-modal Resonance through Evidential Deep Learning for Enhanced Zero-Shot Learning 02/2024-04/2024

- **Role:** Led the entire project, focusing on developing the core methodology for robust multimodal resonance in zero-shot learning.
- **Achievements:** Enhanced the reliability of modal fusion and decision-making, achieving remarkable performance in open-domain zero-shot scenarios, addressing challenges related to data scarcity and uncertainty.

FineCLIPER: Multi-modal Fine-grained CLIP for Dynamic Facial Expression Recognition with AdaptERs 01/2024-03/2024

- **Role:** Managed the project, pioneering fine-grained semantic extraction technology and adapter-based fine-tuning methods.
- **Achievements:** Notably enhanced CLIP's comprehension of motion features and emotional semantics, boosting its capability in dynamic facial expression recognition.

Evidential Deep Partial Multi-View Classification With Discount Fusion

11/2023-01/2024

- **Role:** Oversaw project management, focusing on core methodology. Introduced learnable weighting factors to enhance evidence filtering in multi-view classification.
- **Achievements:** Improved accuracy and robustness, ensuring reliable decisions even with incomplete or conflicting views.

Trusted Unified Feature-Neighborhood Dynamics for Multi-View Classification

09/2023-11/2023

- **Role:** Lead the entire project process, developing core methods for unified feature neighborhood representation and integrated dynamic Markov random field fusion.
- **Achievements:** Achieved superior accuracy, reliability, and stability in complex, high-conflict scenarios, outperforming traditional methods for enhanced decision precision.

Towards Robust Evidential Partial Multi-view Classification

06/2023-08/2023

- **Role:** Provided comprehensive guidance for the project, emphasizing iterative learning to optimize evidence distribution in partial multi-view settings.
- **Achievements:** Significantly improved decision accuracy and reliability under incomplete data, ensuring robust performance through multi-round optimizations.

BoViLA: Bootstrapping Video-Language Alignment via LLM-Based Self-Questioning and Answering

04/2023-06/2023

- **Role:** Managed project, pioneering self-questioning paradigm to enhance video-language alignment.
- **Achievements:** Activated LLM-based multimodal alignment, boosting precision and contextual awareness in video-language tasks via high-quality question generation.

Beyond Uncertainty: Evidential Deep Learning for Robust Video Temporal Grounding 02/2023-04/2023

- **Role:** Led the project, focusing on systematically addressing video temporal uncertainty via novel evidence-based deep learning.
- **Achievements:** Significantly enhanced model robustness and interpretability for open-ended user queries, setting a new standard for uncertainty management in video grounding tasks.

Adaptive Weighted Multi-View Evidential Clustering with Feature Preference 07/2022-04/2023

- **Role:** Led algorithm implementation, experimental design, and report writing, while contributing insights to the project's core concepts and design discussions.
- **Achievements:** Extended traditional evidence clustering to multi-view clustering through a unified optimization framework integrating member and view feature weighting, enhancing accuracy and enabling interpretable and cautious clustering in multi-view scenarios, driven by Dempster-Shafer theory.

A Belief Logarithmic Similarity Measure Based on Dempster-Shafer Theory and Its Application in Multi-Source Data Fusion 12/2022-02/2023

- **Role:** Led conceptual design, code implementation, manuscript writing, and discussions at each stage
- **Achievements:** Resolved Dempster-Shafer (D-S) framework's instability in handling conflicting evidence, validated its effectiveness across multiple tasks, and demonstrated exceptional scalability in areas such as information fusion, zero-shot learning, video analysis, stereo vision, person recognition, and cross-modal retrieval.

PATENT CERTIFICATE

An Autonomous Snow-Sweeping Robot, with Publication Number: CN202121512613 (Sole First Inventor)

Selected Publications

Trusted AI: Focused on exploring the potential of Evidence Theory from a set and distribution perspective for robust multimodal alignment, and tailoring credible AI solutions for downstream tasks such as multi-view and multi-modal applications.

- Adaptive Weighted Multi-View Evidential Clustering With Feature Preference (Student First Author, KBS, Q1 Top)
- L2-Regularization Based Two-Way Weighted Neutrosophic Clustering With Manhattan and Euclidean Distances (Student Second Author, FSS, Q1 Top)
- CREST: Cross-modal Resonance through Evidential Deep Learning for Enhanced Zero-Shot Learning (First Author, ACM MM 2024)
- Evidential Deep Partial Multi-View Classification With Discount Fusion (First Author, arXiv, under review at Information Fusion)
- Trusted Unified Feature-Neighborhood Dynamics for Multi-View Classification (First Author, arXiv, under review at AAAI 2025)
- Towards Robust Evidential Partial Multi-view Classification (Student First Author, arXiv, under review at AAAI 2025)
- REMEDY: Robust Evidential Multimodal Named Entity Discrimination Yielding Deep Insights (First Author, arXiv)
- BoViLA: Bootstrapping Video-Language Alignment via LLM-Based Self-Questioning and Answering (Third Author, arXiv, under review at AAAI 2025)

Personalized AGI: Introduced GaussianVTON, a 3D virtual try-on framework (The first 3D virtual try-on framework based on image prompts) leveraging Gaussian Splatting for image-cue-driven 3D editing; Researched and designed a three-stage refinement strategy, extensively validated through experiments for effectiveness in 3D virtual try-on and 3D editing.

- GaussianVTON: 3D Human Virtual Try-ON via Multi-Stage Gaussian Splatting Editing with Image Prompting (Third Author, arXiv)

Video Understanding: Explored the potential of fine-tuning CLIP for dynamic expression recognition, proposing an uncertainty evaluation technique for video localization, successfully unifying precision, interpretability, and robustness in video understanding.

- FineCLIPER: Multi-modal Fine-grained CLIP for Dynamic Facial Expression Recognition with AdaptERs (Second Author, ACM MM 2024)
- Beyond Uncertainty: Evidential Deep Learning for Robust Video Temporal Grounding (Co-First Author, arXiv, under review at NeurIPS 2024)

AWARDS / HONORS

Competition Wins:

- M Award at the Mathematical Contest in Modeling (MCM) for Undergraduate Students, USA
- Provincial First Prize at the National College Students Mathematical Modeling Contest
- Provincial Second Prize at the Northeast China Mathematical Modeling League

Kaggle Expert:

- Silver Medals (Top 2%)*3 in CV-related Competitions
- Silver Medals (Top 3%)*2 in NLP-related Competitions
- Bronze Medal (Top 7%)*1 in an NLP-related Competition

University-level Honors:

- University-level First-Class Scholarship (once)
- University-level Second-Class Scholarship (twice)
- University-level Third-Class Scholarship (twice)
- Third Place in the "Electromechanical Cup" Pull-up Team Competition

SERVICES

Conference Reviewer

- The International Conference on Learning Representations (ICLR)
- Neural Information Processing Systems (NeurIPS)
- ACM International Conference on Multimedia (ACM MM)

Clubs and Organizations

- CareerSynapse (Leader)

SKILLS

- **IT skills:** Python (PyTorch&TensorFlow), Java, C/C++, JavaScript, CSS, HTML, MATLAB, Linux, MySQL, OpenCV, React-Native, Gurobi, Git, Shell, LaTeX, Markdown, Web Establishment, Qt, SPSS, Arduino
- **Languages:** English (fluent, CET-6, IELTS 6.5); Mandarin (native)