

Yuhui Hong

Luddy School of Informatics, Computing, and Engineering
Indiana University Bloomington
700 N. Woodlawn Avenue
Bloomington, IN 47408

✉ Email: yuhhong@iu.edu
👤 Website: josiehong.github.io
🐙 GitHub: github.com/JosieHong
🔍 Google Scholar: Yuhui Hong

EDUCATION

Indiana University Bloomington
Ph.D. in Computer Science

Bloomington, IN, US
Sep. 2020 –Jul. 2025 (expect)

Xidian University
B.S. in Computer Science and Technology

Xi'an, Shaanxi, China
Sep. 2015–Jul. 2019

RESEARCH INTERESTS

My research explores the intersection of **deep learning**, bioinformatics, and cheminformatics, with a focus on advancing the **identification of small molecules** in two pathways. The first involves predicting tandem mass spectra and other molecular properties from 3D structures, addressing gaps—often referred to as the “dark matter”—in existing spectral reference libraries. The second approach moves beyond the traditional reliance on database-driven methods by predicting compounds directly from tandem mass spectra. My long-term goal is to use computational and machine learning methods to accelerate **the discovery of novel drugs and natural products**. Additionally, I am passionate about developing **reliable and interpretable neural networks** for real-world applications.

RESEARCH EXPERIENCE

Indiana University Bloomington
Research Assistant
Advisor: Prof. Haixu Tang

Bloomington, IN, US
Sep. 2020 –Present

- Developed deep learning and computational methodologies advanced small molecule identification and analysis.
- Designed interpretable neural network architectures for human microbiome data analysis.

The First Affiliated Hospital of Nanchang University
Research Intern
Mentor: Prof. Sujun Li

Nanchang, Jiangxi, China
May 2021 –Jul. 2021

- Developed a machine learning model for MHC binding prediction using BERT.

Xi'an Jiaotong University
Research Assistant
Advisor: Prof. Yaochen Li

Xi'an, Shaanxi, China
Sep. 2019 –Jul. 2020

- Presented a point-based representation method and benchmarked deep learning models for object tracking and segmentation in traffic images and videos.

PUBLICATIONS

Two of the works, 3DMolMS for retention time and collision-cross section [3] and 3DMolCSP for enantioselectivity [2], have been evaluated on internal data and positively considered for application by pharmaceutical and agricultural leaders, including **Merck**, **AbbVie** and **Corteva**.

BOOKS

1. Qingyang Xiao, Kaiyuan Liu, **Yuhui Hong** & Haixu Tang (2024). “Neural Networks for Chemists.” *American Chemical Society*, DOI:10.1021/acsinfocus.7e8012. [Primer]

PEER-REVIEWED ARTICLES

1. **Yuhui Hong**, Yuzhen Ye & Haixu Tang (2024). "Machine Learning in Small-Molecule Mass Spectrometry." *Annual Review of Analytical Chemistry*. (In press, to be published on May 2025)
2. **Yuhui Hong**, Christopher J Welch, Patrick Piras, & Haixu Tang (2024). "Enhanced Structure-Based Prediction of Chiral Stationary Phases for Chromatographic Enantioseparation from 3D Molecular Conformations." *Analytical Chemistry*, 96(6), 2351-2359. [Paper] [Code]
3. **Yuhui Hong**, Sujun Li, Christopher J Welch, Shane Tichy, Yuzhen Ye, & Haixu Tang (2023). "3DMolMS: Prediction of Tandem Mass Spectra from Three Dimensional Molecular Conformations." *Bioinformatics*, btad354. [Paper] [Code] [PyPI package] [Service on GNPS] [Service on Konia]
4. Yifan Zhang, Zhaojie Hu, Xueqiang Wang, **Yuhui Hong**, Yuhong Nan, XiaoFeng Wang, Jiatao Cheng & Luyi Xing (2024). "Navigating the Privacy Compliance Maze: Understanding Risks with Privacy-Configurable Mobile SDKs." In *33rd USENIX Security Symposium*, pp. 6543-6560. [Paper]
5. Yaochen Li, **Yuhui Hong**, Yonghong Song, Chao Zhu, Ying Zhang, & Ruihao Wang (2022). "SiamPolar: Semi-supervised Realtime Video Object Segmentation with Polar Representation." *Neurocomputing*, 467, 491-503. [Paper] [Code]
6. Yaochen Li, Chao Zhu, Yuehu Liu, **Yuhui Hong**, & Jianji Wang (2021). "Geometric and Semantic Analysis of Road Image Sequences for Traffic Scene Construction." *Neurocomputing*, 465, 336-349. [Paper] [Code]

ONGOING ARTICLES

1. **Yuhui Hong**, Sujun Li, Yuzhen Ye, & Haixu Tang (2024). "FIDDLE: a deep learning method for chemical formulas prediction from tandem mass spectra." *bioRxiv*, 2024-11. (Under review). [Preprint] [Code]
2. Mahsa Monshizadeh*, **Yuhui Hong***, & Yuzhen Ye (2024). "Multitask Knowledge-primed Neural Network for Predicting Missing Metadata and Host Phenotype based on Human Microbiome." *bioRxiv*, 2024-02. (Under review). [Preprint] [Code]
3. Ludwig Lautenbacher, Kevin L. Yang, Tobias Kockmann, Christian Panse, Wassim Gabriel, Dulguun Bold, Elias Kahl, Matthew Chambers, Brendan X. MacLean, Kai Li, Fengchao Yu, Brian C. Searle, Wilburn, Damien, Mohammad Reza Zare Shahneh, **Yuhui Hong**, Haixu Tang, Mingxun Wang, Ralf Gabriels, Robbin Bouwmeester, Robbe Devreese, Tobias K. Schmidt, Alexey I. Nesvizhskii, & Mathias Wilhelm (2024). "Koina: Democratizing machine learning for proteomics research." (Under review). [Website] [Code]

*equal contribution as co-first authors

CONFERENCE PRESENTATIONS

1. **Poster presentation.** "Predicting Compositional Fragments of Compounds from Their Tandem Mass Spectra Using Deep Neural Networks" [Poster]
72nd Conference on Mass Spectrometry and Allied Topics. Jun. 2 - 6, 2024. Anaheim, CA.
2. **Poster presentation.** "3DMolMS: Prediction of Tandem Mass Spectra from 3D Molecular Conformations"
Turkey Run Analytical Chemistry Conference 2023. Sep. 29 - 30, 2023. Marshall, IN.
3. **Oral Presentation** "A Machine Learning Model for Chemical Formula Prediction Using Tandem Mass Spectra of Compounds" [Slides]
71st Conference on Mass Spectrometry and Allied Topics. Jun. 4 - 8, 2023. Houston, TX.
4. **Poster Presentation** "Prediction of Molecular Tandem Mass Spectra Using 3-Dimensional Conformers" [Poster]
70th Conference on Mass Spectrometry and Allied Topics. Jun. 5 - 9, 2022. Minneapolis, MN.

TEACHING EXPERIENCE

Instructor

INFO-I529, Machine Learning Bioinformatics

Indiana University Bloomington

Fall 2024

Assistant Instructor

CSCI-D351, Big Data Analytics
Instructor: Prof. Haixu Tang

Indiana University Bloomington

Fall 2024

PROFESSIONAL SERVICES

- **Reviewer:** BMC Genomics, BMC Bioinformatics, Pharmaceutical Research, Beilstein Journal of Organic Chemistry, Chemical Physics Letters
- **Sub-reviewer:** (conferences) RECOMB 2025, ACM BCB 2024, ISMB 2023, RECOMB 2023, RECOMB 2022; (journals) Analytical Chemistry, International Journal of Mass Spectrometry
assisted in reviewing papers under the guidance of Prof. Haixu Tang

PROFESSIONAL AFFILIATIONS

- American Society for Mass Spectrometry (ASMS), Member. 2022 - Present
- NSF Center for Bioanalytic Metrology, Student. 2021 - Present

SCHOLARSHIPS AND AWARDS

- **Special Academic Scholarship of Xi'an Jiao Tong University** 2019
(Top 20% in the students) Academic Administration of Xi'an Jiao Tong University
- **Second-tier Scholarship of Xidian University** 2018
(Top 10% in the students) Academic Administration of Xidian University
- **Meritorious Winner of MCM (Mathematical Contest In Modeling)** 2018
(Top 10% in the 8085 teams) COMAP(the Consortium for Mathematics and Its Application)