Tongxin Wang

tw11@iu.edu | 614-886-9177
714 Blake St Apt D, Indianapolis, IN 46202

EDUCATION

Indiana University Bloomington — Bloomington, IN
PhD Candidate in Computer Science, Minor in Bioinformatics; GPA: 4.00/4.0 — *Aug. 2017 - Jun. 2021 (Expected)*

The Ohio State University — Columbus, OH
PhD Student in Electrical and Computer Engineering; GPA: 4.00/4.0 — *Aug. 2016 - Aug. 2017*

Fudan University — Shanghai, China
Bachelor of Science in Physics; GPA: 3.64/4.0 — *Aug. 2011 - Jul. 2016*

INDUSTY EXPERIENCE

Facebook, Inc. — Menlo Park, CA
Software Engineer Intern (Machine Learning) — *May 2020 - Aug. 2020*

- Developed machine learning models for feed ranking projects to improve the newsfeed experience for Facebook App users.

Tencent AI Lab — Shenzhen, China
Algorithm Research Intern — *May 2019 - Aug. 2019*

- Developed a pipeline to predict microsatellite instability status of patients from H&E histological whole-slide images using deep residual neural networks implemented by Python and PyTorch.
- Co-first-author paper accepted by *IEEE International Symposium on Biomedical Imaging (ISBI) 2020*.

Bosch IoT Research Group — Shanghai, China
Research Assistant Intern — *Jul. 2015 - Sep. 2015*

- Developed a product registration and validation system using physical unclonable functions technologies and error correction algorithms implemented by C++, Matlab, and OpenCV.
- Tested the robustness and feasibility of the aforementioned system on various packaging materials.

RESEARCH EXPERIENCE

Regenstrief Institute, Indiana University — Adviser: Dr. Kun Huang
Research Assistant — *Aug. 2017 - Present*

- **Deep Transfer Learning for Single-cell RNA Sequencing Batch Correction**
  * Developed BERMUDA, an innovative method to integrate multiple single-cell RNA sequencing datasets and remove batch effects based on deep neural networks with maximum mean discrepancy loss using Python and PyTorch.
  * First-author paper accepted by *Genome Biology*.
- **Topological Methods for Analysis of Single-cell RNA Sequencing Data**
  * Developed a novel method to combine topological data analysis techniques with gene co-expression network analysis for single-cell RNA sequencing data visualization using Python.
  * First-author paper accepted by *Pacific Symposium on Biocomputing (PSB) 2019* with oral presentation.
- **Generalized Gene Co-expression Analysis via Subspace Clustering Using Low-rank Representation**
  * Developed a new, generalized gene co-expression network analysis method based on low-rank representation and a quasi-clique detection algorithm using Matlab and R.
  * First-author paper accepted by *BMC Bioinformatics*.

Department of ECE, The Ohio State University — Adviser: Dr. Kun Huang
Research Assistant, University Fellowship — *Apr. 2017 - Aug. 2017*

- **Diagnosing Effusion Fluid Cytology Using Whole Slide Imaging and Multiple Instance Learning**
  * Developed an algorithm to identify patches with malignant cells in effusion fluid cytology whole slide images based on multiple clustered instance learning using Matlab.

TECHNICAL SKILLS

- **Programming Languages:** Python, R, Matlab, Java, SQL
- **Algorithm Design:** Machine Learning, Data Mining, Biomedical Informatics

COURSEWORK & TRAINING

Machine Learning, Deep Learning, Artificial Intelligence, Data Mining, Data Visualization, Database Concepts, Computer Vision, Image Processing, Data Structure, Algorithms, Bioinformatics

PREPRINTS

Travis S Johnson, Y Yu Christina, Zhi Huang, Siwen Xu, **Tongxin Wang**, Chuanpeng Dong, Wei Shao, Mohammed Abu Zaid, Xiaoqing Huang, Yijie Wang *et al*. Diagnostic Evidence GAuge of Single cells (DEGAS): A flexible deep-transfer learning framework for prioritizing cells in relation to disease. *bioRxiv*.

PUBLICATIONS

Tongxin Wang*, Wei Shao*, Zhi Huang, Haixu Tang, Jie Zhang, Zhengming Ding, Kun Huang. MOGONET integrates multi-omics data using graph convolutional networks allowing patient classification and biomarker identification. *Nature Communications* (IF = 12.1), 2021.

Zhi Huang, Zhi Han, **Tongxin Wang**, Wei Shao, Shunian Xiang, Paul Salama, Maher Rizkalla, Kun Huang, and Jie Zhang. TSUNAMI: translational bioinformatics tool suite for network analysis and mining. *Genomics, Proteomics & Bioinformatics*, 2021.

Tongxin Wang, Zhengming Ding, Wei Shao, Haixu Tang, and Kun Huang. Towards fair cross-domain adaptation via generative learning. *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2021.

Rui Cao, Fan Yang, Si-Cong Ma, Li Liu, Yu Zhao, Yan Li, De-Hua Wu, **Tongxin Wang**, Wei-Jia Lu, Wei-Jing Cai *et al*. Development and interpretation of a pathomics-based model for the prediction of microsatellite instability in colorectal cancer. *Theranostics* (IF = 8.6), 2020.

Wei Shao, **Tongxin Wang**, Liang Sun, Tianhan Dong, Zhi Han, Zhi Huang, Jie Zhang, Daoqiang Zhang, and Kun Huang. Multi-task multi-modal learning for joint diagnosis and prognosis of human cancers. *Medical Image Analysis* (IF = 11.1), 2020.

Tongxin Wang*, Weijia Lu*, Fan Yang*, Li Liu, Zhongyi Dong, Weimin Tang, Jia Chang, Wenjing Huan, Kun Huang, and Jianhua Yao. Microsatellite instability prediction of uterine corpus endometrial carcinoma based on H&E histology whole-slide imaging. *IEEE International Symposium on Biomedical Imaging (ISBI)*, 2020.

Tongxin Wang, Travis S Johnson, Wei Shao, Zixiao Lu, Bryan R Helm, Jie Zhang, and Kun Huang. BERMUDA: a novel deep transfer learning method for single-cell RNA sequencing batch correction reveals hidden high-resolution cellular subtypes. *Genome Biology* (IF = 14.0), 2019.

Tongxin Wang, Jie Zhang, and Kun Huang. Generalized gene co-expression analysis via subspace clustering using low-rank representation. *BMC Bioinformatics*, 2019.

Tongxin Wang, Travis Johnson, Jie Zhang, and Kun Huang. Topological methods for visualization and analysis of high dimensional single-cell RNA sequencing data. *Pacific Symposium on Biocomputing (PSB)*, 2019.

Travis S Johnson, **Tongxin Wang**, Zhi Huang, Christina Y Yu, Yi Wu, Yatong Han, Yan Zhang, Kun Huang, and Jie Zhang. LAmbDA: label ambiguous domain adaptation dataset integration reduces batch effects and improves subtype detection. *Bioinformatics*, 2019.

Wei Shao, **Tongxin Wang**, Zhi Huang, Jun Cheng, Zhi Han, Daoqiang Zhang, and Kun Huang. Diagnosis-guided multi-modal feature selection for prognosis prediction of lung squamous cell carcinoma. *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 2019.

Wei Shao, Zhi Han, Jun Cheng, Liang Cheng, **Tongxin Wang**, Liang Sun, Zixiao Lu, Jie Zhang, Daoqiang Zhang, and Kun Huang. Integrative analysis of pathological images and multi-dimensional genomic data for early-stage cancer prognosis. *IEEE Transactions on Medical Imaging*, 2019.

(* denotes equal contribution)