TONGAN CAI

College of Information Sciences and Technology
Pennsylvania State University
(734) 239-0723 | cta@psu.edu | johncta@umich.edu | c-t-a.me

EDUCATION

The Pennsylvania State University *GPA:3.9/4.0* -Anticipated Dissertation: 10/2024 State College, PA
Ph.D. in Informatics Advisor: James Z. Wang, Distinguished Professor 08/2019 – Present
University of Michigan *GPA:3.7/4.0* -Magna Cum Laude Ann Arbor, MI
Bachelor of Science and Engineering in Data Science (EECS) 08/2017 – 05/2019
Shanghai Jiao Tong University *GPA:3.6/4.0* -Outstanding Graduate-Class of 2019 Shanghai, China
Bachelor of Science and Engineering in Electrical & Computer Engineering 08/2015 – 08/2019

EXPERTISE & INTEREST

7 years of research and work expertise in state-of-the-art Computer Vision, Data Science, Deep Learning, with their applications in healthcare, finances, transportation, and education. Strong passion about techniques for data to "*talk*" themselves, for ML models to improve *fairness* and *usability*, and for AI products to make *positive* impacts.

RESEARCH & WORKING EXPERIENCE

The Pennsylvania State University – College of Information Sciences & Technology University Park, PA
Research Lead – *DeepStroke* Project 08/2019 – Present

- Advance ER stroke triage with mobile AI. Develop multi-sensor systems using speech audio, text transcripts, and facial motion videos to construct efficient, novel multi-modal deep learning frameworks. Participate in proposal writing (multiple NIH R01 at million scale).
- Develop 2D/3D face representation, multi-modal transformers, sequence models, adversarial training, NLP models, and Gen AI for stroke diagnosis. Implement real-time face detection, tracking, and estimation pipelines in OpenCV and Dlib, audio processing with librosa, and diagnosis models in PyTorch/timm.
- Develop an RGB-D face video/speech audio collection tool and a full processing/diagnosis tool on the iOS platform with < 2-minute onboard runtime while outperforming clinical triage by 17% sensitivity and specificity gains, surpassing prior models with 8% in AUC, achieving better fairness.
- Leading top-tier publications: MICCAI 2020 Conference paper on multimodal deep learning with facial image and speech text, Medical Image Analysis 2022 Journal paper on multi-level patient video/audio feature fusion, fairness-aware pre-training, and adversarial training. Two papers in submission, with one on Gen AI/Video retargeting for privacy, and one on mobile AI, state space model, and transformers.

✧ **Key Skills:** Multimodal Deep Learning, 2D/3D Computer Vision, Mobile Dev, NLP, Face Video Processing, Speech Recognition

Nationwide Mutual Insurance Company – Enterprise Analytics Office Columbus, OH/Remote
Data Scientist Intern 05/2023 – 08/2023, 05/2024 – Present

- Conduct exploratory NLP research for Agri-Business's refined cause of loss Analysis. Design a local-hosted RAG pipeline of LangChain and Llama-3 on policy documents for claim tagging and coverage validation. Performed AWS-S3 database linkage with 6M~ unstructured claim and 27.7M~ raw notes. Conduct NER with SpaCy and text clustering with LDA features on claim/policy text documents to extract insights. Visualize with Streamlit App and Bokeh plot.
- Prototype PoC framework for computer vision-based automated building underwriting. A multi-sourced building dataset with ~200 online images is collected and accompanied by labeled manual object bounding boxes using LabelMe. Finetuned YOLOv8 detection model with 90%+ multi-class accuracy. The Segment Anything model (SAM) is deployed as ONNX runtime to provide masks within minutes on CPU, and a multi-threading GUI interface is fabricated in Python.
- Lead team of data scientists in presenting an exploratory data analysis (EDA) on a corporate vehicle policy dataset. Draw business insights from geological analysis of vehicle characteristics and premium information.

✧ **Key Skills:** LLM/RAG, NLP, Text Processing, AWS, Object Detection/Segmentation, Foundation Model, EDA

Tencent America – Medical AI Lab Palo Alto, CA/Remote
Research Scientist Intern 05/2021 – 08/2021

- Improve 3D-based document image de-warping. Synthetic dataset rendered with real document warping meshes and PDF images, cascaded auto encoder networks built for shape, texture, and refinement.
- Design an OCR detection loss with BERT embedding to penalize local distortions and benefit subsequent OCR. Validate model performance improvements in model performance on the benchmarking doc3D dataset.

✧ **Key Skills:** 3D Geometry and Rendering, Computer Vision, Language-guided Image Processing, OCR

The Pennsylvania State University – College of Information Sciences & Technology University Park, PA
Research Assistant – *AI-PLAX* Project 03/2023 – 11/2023

- Design and develop an efficient multi-task vision-language contrastive learning framework for placental images.
- Propose a two-stage text feature re-composition for medical information preservation and data augmentation.
- The proposed method outperforms strong baselines on diagnostic tasks while maintaining computationally efficient.

✧ **Key Skills:** Vision-Language Contrastive Learning, Multi-Task Learning, Digital Pathology

The Pennsylvania State University – College of Information Sciences & Technology University Park, PA
Vision System Data Engineer 05/2020 – 03/2024

- Curate a large video dataset (BoLD) for Human Bodily Expression of Emotion Project with 13k instances and 100k+ frames for public access and AMT labeling. Collaborate with external institutes for large-scale Laban Movement Analysis with in-the-wild movie video data.
- Lead the collection of 3D human action videos/audio with multi-camera Kinect/VICON system. Build, maintain, refine, and debug the GUI/API interfaces for collaborators' use.

✧ **Key Skills:** Data Curation, GUI Programming, Crowd-Sourcing

The Pennsylvania State University – Applied Cognitive Science Lab University Park, PA
Research Project Lead 11/2019 – 03/2021

- Design a quantitative study on bibliography managers (Mendeley, Zotero, EndNote, RefWorks) by evaluating the amount of physical and mental effort users make. Keystrokes and mouse moves were recorded and analyzed.
- Pilot prospective-retrospective survey with tasks to search paper and construct bibliography, revealing preference changes and usability issues. First-authored manuscript received Best Paper in AsianCHI Symposium 2021.

✧ **Key Skills:** User Study, Behavior Analysis, Cognitive Analysis, HCI

University of Michigan - Michigan integrated Center for Health Analytics & Medical Prediction Ann Arbor, MI
Researcher - MiCHAMP 03/2018 – 05/2019

- Build Python pipelines and adopted statistical machine learning models (XGBoost, GBDT, LightGBM, Random Forest, Linear and Logistic Regression) for medical data (HALT-C, NHANES, and MEPS) EDA and modeling.
- Pilot *DocDollars* Nationwide Salary Survey to reveal the demographics-related salary discrepancies for academic physicians. Distribute questionnaires and collect data with RedCap, analyze and visualize data with Python and R.
- Geologically analyze the Hepatitis C situation in Michigan by contrasting the locations of doctors with individuals to identify potential barriers to Hepatitis C medication allocation and delivery.

✧ **Key Skills:** Data Science, Statistical Analysis, Machine Learning, Survey Study, Geospatial Analysis

Zhejiang University of Finance & Economics – School of Information Hangzhou, Zhejiang, China
Machine Learning Researcher – Anomaly Detection for Credit Scoring Project 05/2018 – 07/2018

- Develop an imbalance learning (SMOTE) stacking ensemble method with Bayesian Optimization parameter tuning for credit scoring dataset anomaly detection. Validate on a Breast Cancer Dataset and lead a journal publication.

✧ **Key Skills:** Anomaly Detection, Imbalance Learning, Optimization, Credit Scoring

PATENT

- Wang, J.Z., Yu, M., **Cai, T.**, Huang, X., Wong, K., Volpi, J., Wong, S.T.C.. Systems and methods for assisting with stroke and other neurological condition diagnosis using multimodal deep learning. Pub. No. US 2023/0363679 A1. Filed Sept. 17, 2021. Published Nov. 16, 2023. Patent Pending.

PUBLICATIONS (* INDICATES EQUAL CONTRIBUTIONS)

- **Cai, T.***, Ni, H.*, Ma, Q., Xue, Y., Wong, K., Huang, X., Wang, J.Z., Volpi, J., Wong, S.T.C.. (2024). *SafeScreen*: Facial Video De-identification for AI-Assisted Stroke Screening. Submitted to *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI'24)*.

- **Cai, T.**, Wong, K., Wang, J.Z., Huang, X., Yu, X., Volpi, J., Wong, S.T.C.. (2024). M^3Stroke: Advancing Stroke Triage with **M**ulti**M**odal **M**obile AI for Accurate and Timely Identification of Mild to Moderate Acute Strokes in Emergency Care. *In journal submission.*

- Pan, Y., **Cai, T.**, Metha, M., Gernand, A.D., Goldstein, J.A., Mithal, L., Mwinyelle, D., Gallagher, K., Wang, J.Z.. (2023). Enhancing Automatic Placenta Analysis through Distributional Feature Re-composition in Vision-Language Contrastive Learning. In *International Conference on Medical Image Computing and Computer Assisted Intervention – MICCAI 2023.* pp. 116-126.

- **Cai, T.***, Ni, H.*, Yu, M., Huang, X., Wong, K., Volpi, J., Wang, J.Z., Wong, S.T.C.. (2022). *DeepStroke*: An Efficient Stroke Screening Framework for Emergency Rooms with Multimodal Adversarial Deep Learning. *Medical Image Analysis,* Vol. 80. p. 102522. (IF=13.8)

- **Cai, T.,** Chen, C., Huang, T. H., Ritter, F. E. (2021). What Makes a Good Reference Manager? A Quantitative Analysis of Bibliography Management Applications. In *Asian CHI Symposium 2021*. pp. 64-69. (Best paper)

- Yu, L.*, **Cai**, **T.***. (2021). Ensemble learning for early identification of students at risk from online learning platforms. *Advances in Data Science & Information Engineering*, Springer. pp. 531-542.

- Yu, M.*, **Cai, T.***, Huang, X., Wong, K., Volpi, J., Wang, J.Z., Wong, S.T.C.. (2020). Toward Rapid Stroke Diagnosis with Multimodal Deep Learning. In *International Conference on Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*. pp. 616–626. (Oral)

- Golbus, J. R.*, **Cai**, **T.*,** Najarian, D., Trumpower, B., Kao,T., Waljee, A. K., Nallamothu, B. K.. (2019). Determinants of Compensation for US Academic Physicians: Does Gender Matter? *Circulation: Cardiovascular Quality and Outcomes,* Vol. 12. pp. A225.

- **Cai**, **T.,** He, H., Zhang, W.. (2018). Breast Cancer Diagnosis Using Imbalanced Learning and Ensemble Method. *Applied and Computational Mathematics*. Vol. 7, No. 3. pp. 146-154.

SELECTED AWARDS AND HONORS

• AsianCHI 2021 Best Paper Award	2021
• Outstanding Graduate – Class of 2019 – Shanghai Jiao Tong University	2019
• Undergrads Excellence Scholarship - Shanghai Jiao Tong University	2018, 2017, 2016
• Dean's List - University of Michigan	2019,2018,2017
• Dean's List - Shanghai Jiao Tong University	2017, 2016, 2015

TECHNICAL SKILLS

Domain Expertise: Computer Vision (Medical Image, Object Detection/ Recognition, Segmentation, Motion Recognition and Tracking, Multimodal, 3D Reconstruction, Contrastive Learning, Generative Models), Natural Language Processing (Sentiment Analysis, Speech Recognition, LLM/RAG, NER), Data Science (Visualization, Modeling, Management, Map-Reduce, A/B Testing)

Coding: Python, R, HTML/CSS, MySQL, C++/C, Java, MATLAB, SAS, Swift C#, Shell/Bash, LaTeX, JavaScript, D3

Packages: PyTorch, TensorFlow, Keras, PIL, OpenCV, Dlib, Transformers, scikit-learn, Matplotlib, NLTK, Gensim

Platform: AWS (SageMaker, S3, EC2), Git, CUDA, JIRA, Slurm, Hugging Face, Linux, Windows, Mac OS

Technical: OOP, Web crawling/develop, Android/iOS develop, Math/Scientific Programming, MLOps, CI/CD

ADDITIONAL INFORMATION

Professional Service

- Invited reviewer for npj Digital Medicine, *Springer* (IF=15.2)
- Invited reviewer for Medical Image Analysis, *Elsevier* (IF=13.8)
- Invited reviewer for IEEE Journal of Biomedical and Health Informatics (JBHI), *IEEE* (IF=7.7)
- Invited reviewer for Artificial Intelligence in Medicine (AIIM), *Elsevier* (IF=7.5)
- Invited reviewer for Computerized Medical Imaging and Graphics (CMIG), *Elsevier* (IF=5.7)
- Invited reviewer for Innovation and Research in BioMedical engineering (IRBM), *Elsevier* (IF=4.8)
- Invited reviewer for Computer Vision and Image Understanding (CVIU), *Elsevier* (IF=4.5)
- Invited reviewer for IEEE Access, *IEEE* (IF=3.9)
- Invited reviewer for The Journal of Supercomputing, *Springer*
- Invited reviewer for Quantitative Imaging in Medicine and Surgery (QIMS)
- Invited reviewer for Computational Intelligence and Neuroscience, *Hindawi*
- Reviewer for European Conference on Computer Vision (ECCV)
- Reviewer for Medical Image Computing and Computer Assisted Intervention Conference (MICCAI)

Certificates

- Social and Behavioral Human Subjects Research (IRB) Certified – Penn State University.