

# Youbao Tang

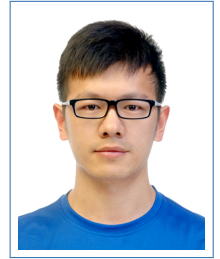
STAFF RESEARCHER

PAII Inc., USA

☎ (1) (240)899-8968

✉ [tybxiaobao@gmail.com](mailto:tybxiaobao@gmail.com)

📄 <https://tangyoubao.github.io/>



## Summary

- Accomplished Research Scientist with 10+ years of experience in developing cutting-edge solutions in multi-modal large language models, human-centered video analysis, medical imaging analysis, and computer vision.
- Spearheaded the deployment of AI-driven systems for diverse applications, including a multi-modal medical consultation assistant achieving 98% response acceptance rate, sound/music-driven motion generation, image and video generation, and intelligent PACS tools for automated CT scan analysis.
- Recognized with multiple prestigious awards, including the RSNA Trainee Research Prize and NIH Fellows Award of Research Excellence, 15+ worldwide patents, and 50+ research papers (CVPR, ECCV, ACMMM, MICCAI, IEEE TIP, IEEE TMI, MIA, etc).
- Adept at bridging research and practical deployment, with a passion for delivering impactful AI innovations in healthcare, finance, multimedia, and computer vision applications.

## Experience

**PAII Inc. Staff Research Scientist, Tech Leader, PAII Inc., Palo Alto, CA, USA (Remote).**

Jan. 2023  
–Present

- **Multi-modal Large Language Model (MLLM):**
  - 1) I led an ongoing project of developing an end-to-end speech interaction MLLM, which can understand multiple modalities (including text, image, audio, and video) and output audio responses with low latency. It will be further fine-tuned and deployed in specific domain scenarios to provide low-latency voice conversational chatbot services.
  - 2) I led a research project on MLLM for healthcare AI assistants, designed to facilitate effective communication with online patients using their text and image inputs. The AI assistant efficiently supports doctors by gathering relevant patient information. The model was developed to accurately extract features from images, comprehend historical context and input queries, and deliver precise, context-aware responses.
  - 3) Traditional online medical consultation systems require many professional doctors as patient volumes increase, leading to high operational costs. To address this challenge, we developed an MLLM-based AI assistant tailored for multiple specialties, such as dermatology and pediatrics. This assistant has been successfully deployed on an online consultation platform, achieving an impressive response acceptance rate of over 98% in real-world scenarios.
- **Motion and Video Generation:**
  - 1) I led an open-source-based project leveraging SDXL, DreamBooth, and StoryDiffusion to bring text-based stories to life through consistent images or videos. This solution was successfully deployed for credit card promotion, showcasing a comic series featuring the credit card mascot across multimedia social platforms.
  - 2) Traditional digital and visual human applications often rely on predefined, fixed characters and motions, making the addition of new features costly and inefficient. To address this, we developed a sound/music-driven motion generation approach and a realistic human video generation method utilizing motion and reference transfer. The motion generation approach creates smooth, sound-aligned motions, while the video generation method produces high-quality human videos with seamless content consistency.

**PAII Inc. Senior Research Scientist II, PAII Inc., Palo Alto, CA, USA (Remote).**

Jul. 2020  
–Dec. 2022

- **Human-centered Video Analysis:** Driven by the business requirements, I built a human-centered video analysis framework, including three main parts: pedestrian detection and tracking, 2D/3D pose estimation and mesh recovery, and action localization and recognition. Besides using existing techniques, we also developed several approaches to address the issues left by them. It was successfully deployed to automate some bank-related businesses, e.g., people flow analysis, employee misconduct detection, dual-recording, and video & audio-based speech recognition for liveness detection.
- **Universal Lesion Analysis in CT Images:** In this project, we defined and produced a universal lesion analysis product that can detect different kinds of lesions across various organs in the chest and abdomen on CT scans. As a groundbreaking product, it can be used for patients' comprehensive screening. We collaborated with several hospitals, achieved promising results on real clinical data, and received positive feedback from doctors.
- **Intelligent PACS Development:** In this project, we developed an intelligent PACS product to automate several processing steps (anatomical point matching, lesion tracking, measuring and segmenting) when doctors read CT scans, which are time-consuming and tedious. We have deployed this product to several hospitals and have been working with several top PACS companies in China, to expand its clinical impacts and commercialize it.

- NIH Postdoctoral Research Fellow, National Institutes of Health (NIH), Bethesda, MD, USA.**
- Jun. 2017 – Jun. 2020
- **Lesion Detection, Segmentation, and RECIST Diameter Prediction on CT Scans:** Several deep learning-based models were developed to tackle three important clinical applications of medical image analysis: universal lesion detection, segmentation and RECIST (response evaluation criteria in solid tumors) diameter prediction on CT scans. One work won the **2020 RSNA Trainee Research Prize**, the best paper in the informatics category. Another one won the **2019 NIH Fellows Award of Research Excellence**.
  - **Disease Analysis in Chest X-Ray Images:** A series of deep learning-based methods were proposed for different computer-aided diagnosis tasks in Chest X-Ray (CXR) images including disease classification and localization, CXR interpretation, and abnormal lung segmentation. They can handle the situation of lacking annotations of abnormal CXRs for training and deal with the domain shift problem. One work (second author) won the **2019 RSNA Trainee Research Prize**.
- HIT Ph.D., Harbin Institute of Technology (HIT), Harbin, Heilongjiang, China.**
- Aug. 2011 – Dec. 2016
- **Salient Object Detection and Segmentation:** Five approaches were proposed in this project, where one extracted hand-crafted features and employed graph-structural agglomerative clustering to detect the salient objects coarse-to-fine, and the other four were developed based on deep convolutional neural networks (CNN) and adversarial learning.
  - **Scene Text Detection in Natural Images:** Four approaches were proposed in this project, where I used the text edge information, superpixel-based stroke feature transforms, and text-aware saliency detection to generate candidate text regions. Some hand-crafted low-level features and CNN-based deep high-level features were extracted and fused for region classification.
  - **Offline Text-Independent Writer Identification:** Four approaches were proposed in this project, which focused on the extraction of discriminative features (including hand-crafted features and CNN-based features) identified in images of handwriting to represent identities and writing habits.
  - **Signature Video Based Personal Recognition:** I drafted the project application. To verify its feasibility and effectiveness, I developed a tool for data collection, collected signature videos from over 60 people, applied image processing techniques to detect, segment, and track the hands and signatures, and extracted some behavior and identity-related features for identification.
- HIT Master, Harbin Institute of Technology (HIT), Harbin, Heilongjiang, China.**
- Aug. 2009 – Jul. 2011
- **Contactless Identity Recognition by Integrating Hand Vein, Handprint, and Handshape:** I was in charge of the algorithm implementation and optimization and the whole system development. Also, I built an Android-based palmprint identification App, which had been used for student checking-in in Prof. Wu's class. In this project, we comprehensively extracted features from hands for identification and verification that can represent our unique personal identity information.
  - **Data Encryption and Decryption Based on Palmprint Features:** I was in charge of encryption and decryption algorithm research and system implementation. In this project, I generated long, complex, and irreversible sequences by encoding the features extracted from palmprint images to encrypt and decrypt data.

## Education

- Sep. 2011 – Dec. 2016
- Doctor of Computer Applied Technology, Harbin Institute of Technology (HIT), China, Recommendation.**  
 Supervisor: Prof. Xiangqian Wu  
 Thesis: Research on Visual Saliency Detection Method and Its Application
- Sep. 2009 – Jul. 2011
- Master of Computer Science and Technology, Harbin Institute of Technology (HIT), China, Recommendation.**  
 Supervisor: Prof. Xiangqian Wu  
 Thesis: Research on Encryption and Decryption Algorithm based on Non-contact Palmprint Feature
- Sep. 2005 – Jul. 2009
- Bachelor of Bioinformatics, Harbin Institute of Technology (HIT), China, Top 20%.**  
 Supervisor: Prof. Xiangqian Wu  
 Thesis: Keystroke-based User Identification and Verification

## Selected Publications

Google scholar citations: 2600; h-index: 29. (2024-11-01)

★ indicates equal contribution, interns under my mentoring.

### JOURNALS

- 8 K. Yan, J.Z. Cai, D.K. Jin, S. Miao, A. Harrison, D.Z. Guo, **Y.B. Tang**, J. Xiao, J.J. Lu, L. Lu, "SAM: Self-supervised Learning of Pixel-wise Anatomical Embeddings in Radiological Images", *IEEE Transactions on Medical Imaging*, 41(10):2658-2669, 2022.
- 7 **Y.B. Tang**, Y.X. Tang, Y.Y. Zhu, J. Xiao, R. Summers, "A Disentangled Generative Model for Disease Decomposition in Chest X-rays via Normal Image Synthesis", *Medical Image Analysis*, 67:101839, 2021.
- 6 Y.X. Tang, **Y.B. Tang**, Y.F. Peng, K. Yan, M. Bagheri, B. Redd, C. Brandon, Z.Y. Lu, M. Han, J. Xiao, R. Summers, "Automated Abnormality Classification of Chest Radiographs Using Deep Convolutional Neural Networks", *NPJ Digital Medicine*, 3(1):1-8, 2020.

- 5 K. Yan, J.Z. Cai, Y.J. Zheng, A. Harrison, D.K. Jin, **Y.B. Tang**, Y.X. Tang, L.Y. Huang, J. Xiao, L. Lu, "Learning from Multiple Datasets with Heterogeneous and Partial Labels for Universal Lesion Detection in CT", *IEEE Transactions on Medical Imaging*, 40(10):2759-2770, 2020.
- 4 **Y.B. Tang**, X.Q. Wu, "Salient Object Detection Using Cascaded Convolutional Neural Networks and Adversarial Learning", *IEEE Transactions on Multimedia*, 21(9):2237-2247, 2019.
- 3 **Y.B. Tang**, X.Q. Wu, "Scene Text Detection using Superpixel based Stroke Feature Transform and Deep Learning based Region Classification", *IEEE Transactions on Multimedia*, 20(9):2276-2288, 2018.
- 2 **Y.B. Tang**, X.Q. Wu, "Scene Text Detection and Segmentation via Cascaded Convolution Neural Networks", *IEEE Transactions on Image Processing*, 26(3):1509-1520, 2017.
- 1 X.Q. Wu (Supervisor), **Y.B. Tang**, Wei Bu, "Offline Text-Independent Writer Identification Based on Scale Invariant Feature Transform", *IEEE Transactions on Information Forensics and Security*, 9(3):526-536, 2014.

## CONFERENCES

- 22 C.Y. Zhang, **Y.B. Tang**, N. Zhang, R.S. Lin, M. Han, J. Xiao, S. Wang, "Bidirectional Autoregressive Diffusion Model for Dance Generation", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.
- 21 V. Agarwal, M. Levy, M. Ehrlich, **Y.B. Tang**, N. Zhang, A. Shrivastava, "Coarse to Fine Human Mesh Recovery with Transformers", *T-CAP Workshop at ECCV*, 2024.
- 20 Y.F. Liu, **Y.B. Tang**, N. Zhang, R.S. Lin, H.Q. Wang, "Prior-enhanced Temporal Action Localization using Subject-aware Spatial Attention", *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) [Oral]*, 2023.
- 19 **Y.B. Tang**, N. Zhang, Y.R. Wang, S.H. He, M. Han, J. Xiao, R.S. Lin, "Accurate and Robust Lesion RECIST Diameter Prediction and Segmentation with Transformers", *International Conference on Medical Image Computing & Computer Assisted Intervention (MICCAI)*, 2022.
- 18 **Y.B. Tang**, J.Z. Cai, K. Yan, L.Y. Huang, G.T. Xie, J. Xiao, J.J. Lu, G.G. Lin, L. Lu, "Weakly-Supervised Universal Lesion Segmentation with Regional Level Set Loss", *International Conference on Medical Image Computing & Computer Assisted Intervention (MICCAI)*, 2021. [**Early accept**]
- 17 **Y.B. Tang**, K. Yan, J.Z. Cai, L.Y. Huang, G.T. Xie, J. Xiao, J.J. Lu, G.G. Lin, L. Lu, "Lesion Segmentation and RECIST Diameter Prediction via Click-driven Attention and Dual-path Connection", *International Conference on Medical Image Computing & Computer Assisted Intervention (MICCAI)*, 2021.
- 16 J.N. Chen, K. Yan, Y.D. Zhang, **Y.B. Tang**, X. Xu, S.W. Sun, Q.P. Liu, L.Y. Huang, J. Xiao, A. Yuille, Y. Zhang, L. Lu, "Sequential Learning on Liver Tumor Boundary Semantics and Prognostic Biomarker Mining", *International Conference on Medical Image Computing & Computer Assisted Intervention (MICCAI) [Student Travel Award]*, 2021.
- 15 J.Z. Cai, **Y.B. Tang**, K. Yan, A. Harrison, J. Xiao, Gigin Lin, L. Lu, "Deep Lesion Tracker: Monitoring Lesions in 4D Longitudinal Imaging Studies", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.
- 14 **Y.B. Tang**, Y.X. Tang, Y.Y. Zhu, J. Xiao, R. Summers, "E<sup>2</sup>Net: An Edge Enhanced Network for Accurate Liver and Tumor Segmentation on CT Scans", *International Conference on Medical Image Computing & Computer Assisted Intervention (MICCAI)*, 2020. [**Early accept**]
- 13 **Y.B. Tang**, K. Yan, J. Xiao, R. Summers, "One Click Lesion RECIST Measurement and Segmentation on CT Scans", *International Conference on Medical Image Computing & Computer Assisted Intervention (MICCAI)*, 2020. [**Early accept**]
- 12 Y.Y. Zhu, **Y.B. Tang**, Y.X. Tang, D. Elton, S. Lee, P. Pickhardt, R. Summers, "Cross-Domain Image Translation by Shared Latent Gaussian Mixture Model", *International Conference on Medical Image Computing & Computer Assisted Intervention (MICCAI)*, 2020.
- 11 Y.X. Tang, **Y.B. Tang**, V. Sanfort, J. Xiao, R. Summers, "TUNA-Net: Task-oriented UNsupervised Adversarial Network for Disease Recognition in Cross-Domain Chest X-rays", *International Conference on Medical Image Computing & Computer Assisted Intervention (MICCAI)*, 2019.
- 10 K. Yan, **Y.B. Tang**, Y.F. Peng, V. Sanfort, M. Bagheri, Z.Y. Lu, R. Summers, "MULAN: Multitask Universal Lesion Analysis Network for Joint Lesion Detection, Tagging, and Segmentation", *International Conference on Medical Image Computing & Computer Assisted Intervention (MICCAI)*, 2019.
- 9 **Y.B. Tang**<sup>\*</sup>, Y.X. Tang<sup>\*</sup>, J. Xiao, R. Summers, "XLSor: A Robust and Accurate Lung Segmentor on Chest X-Rays Using Criss-Cross Attention and Customized Radiorealistic Abnormalities Generation", *International Conference on Medical Imaging with Deep Learning (MIDL) [Spotlight]*, 2019.
- 8 **Y.B. Tang**<sup>\*</sup>, J.Z. Cai<sup>\*</sup>, L. Lu, A. Harrison, K. Yan, J. Xiao, L. Yang, R. Summers, "CT Image Enhancement Using Stacked Generative Adversarial Networks and Transfer Learning for Lesion Segmentation Improvement", *International Conference on Machine Learning in Medical Imaging (MICCAI-MLMI) [Oral]*, 2018.

- 7 **Y.B. Tang**, A. Harrison, M. Bagheri, J. Xiao, R. Summers, "Semi-Automatic RECIST Labeling on CT Scans with Cascaded Convolutional Neural Networks", *International Conference on Medical Image Computing & Computer Assisted Intervention (MICCAI)*, 2018.
- 6 J.Z. Cai\*, **Y.B. Tang\***, L. Lu, A. Harrison, K. Yan, J. Xiao, L. Yang, R. Summers, "Accurate Weakly-Supervised Deep Lesion Segmentation using Large-Scale Clinical Annotations: Slice-Propagated 3D Mask Generation from 2D RECIST", *International Conference on Medical Image Computing & Computer Assisted Intervention (MICCAI)*, 2018.
- 5 D.K. Jin, Z.Y. Xu, **Y.B. Tang**, A. Harrison, D. Mollura, "CT-Realistic Lung Nodule Simulation from 3D Conditional Generative Adversarial Networks for Robust Lung Segmentation", *International Conference on Medical Image Computing & Computer Assisted Intervention (MICCAI)*, 2018.
- 4 **Y.B. Tang**, X.Q. Wu, "Salient Object Detection with Chained Multi-scale Fully Convolutional Network", *ACM Multimedia Conference (ACMMM)*, 2017.
- 3 **Y.B. Tang**, X.Q. Wu, "Saliency Detection via Combining Region-level and Pixel-level Predictions with CNNs", *European Conference on Computer Vision (ECCV)*, 2016.
- 2 **Y.B. Tang**, X.Q. Wu, W. Bu "Deeply-Supervised Recurrent Convolutional Neural Network for Saliency Detection", *ACM Multimedia Conference (ACMMM)*, 2016.
- 1 **Y.B. Tang**, X.Q. Wu, W. Bu, "Saliency Detection Based on Graph-Structural Agglomerative Clustering", *ACM Multimedia Conference (ACMMM)*, 2015.

### CLINICAL ABSTRACTS

- 11 **Y.B. Tang**, et al., "A Unified Framework for Accurate and Robust Lesion RECIST Diameter Prediction and Segmentation with Transformers", *The Radiological Society of North America (RSNA)*, 2022.
- 10 **Y.B. Tang**, et al., "Accurate Weakly-supervised Volumetric Universal Lesion Segmentation Using Large-scale Clinical RECIST Diameter Annotations And Regional Level Set Loss", *The Radiological Society of North America (RSNA)* [Oral], 2021.
- 9 **Y.B. Tang**, et al., "Automatic RECIST Measurement In Longitudinal CT Imaging Studies", *The Radiological Society of North America (RSNA)* [Oral], 2021.
- 8 **Y.B. Tang**, et al., "Automatically, Precisely, And Comprehensively Measuring Tumor Sizes With Minimal Human Effort", *The Radiological Society of North America (RSNA)*, 2021.
- 7 **Y.B. Tang**, et al., "One Click Guided Automatic RECIST Lesion Measurement and Segmentation on CT Scans", *The Radiological Society of North America (RSNA)* [Oral & Featured Paper & RSNA Trainee Research Prize], 2020.
- 6 **Y.B. Tang**, et al., "Accurate Liver and Tumor Segmentation on CT Scans Using an Edge-enhanced Network", *The Radiological Society of North America (RSNA)* [Oral], 2020.
- 5 **Y.B. Tang**, et al., "Making RECIST Measurements Easy: A Semi-Automated Deep Learning System With Expert-Equivalent Accuracy and Better Consistency", *The Radiological Society of North America (RSNA)* [Oral], 2020.
- 4 **Y.B. Tang**, et al., "An Interpretable Generative Model for Chest X-ray Decomposition via Synthesizing Radio-realistic Normal Chest X-rays and Separating Abnormalities", *The Radiological Society of North America (RSNA)*, 2019.
- 3 Y.X. Tang, **Y.B. Tang**, et al., "Abnormal Chest X-ray Identification with Generative Adversarial One-Class Classifier", *The Radiological Society of North America (RSNA)* [Oral & RSNA Trainee Research Prize], 2019.
- 2 **Y.B. Tang**, et al., "Semi-Automatic RECIST Labeling on CT Scans with Cascaded Convolutional Neural Networks", *The Radiological Society of North America (RSNA)*, 2018.
- 1 **Y.B. Tang\***, J.Z. Cai\*, et al., "CT Image Enhancement Using Stacked Generative Adversarial Networks and Transfer Learning for Lesion Segmentation Improvement", *The Radiological Society of North America (RSNA)*, 2018.

### Awards

- 2022 Business Contribution Award (Project Award) of the Group's Science and Technology Development Committee
- 2022 Top Ten Outstanding Projects of the Group's Science and Technology Development Committee
- 2020 RSNA 2020 Trainee Research Prize. (First author)
- 2019 RSNA 2019 Trainee Research Prize. (Second author)
- 2018 NIH Fellows Award of Research Excellence (FARE) 2019 competition.
- 2014 National Scholarship



## Professional Services

### REVIEWER

IEEE Transactions on Pattern Analysis and Machine Intelligence

IEEE Transactions on Information Forensics and Security

IEEE Transactions on Neural Networks and Learning Systems

IEEE Transactions on Industrial Informatics

IEEE Transactions on Image Processing

IEEE Transactions on Medical Imaging

IEEE Transactions on Multimedia

IEEE Transactions on Circuits and Systems for Video Technology

IEEE Transactions on Very Large Scale Integration Systems

IEEE Transactions on Dependable and Secure Computing

IEEE Journal of Biomedical and Health Informatics

Medical Image Analysis

CVPR (2021 Outstanding Reviewer), ICLR, NeurIPS, ICCV, ECCV, AAAI, IJCAI, MICCAI, etc.

### TALKS

- 8 Universal Lesion Analysis in CT Images and Its Application in Intelligent PACS. *Synced AI*. Virtual, 2021-05-23.
- 7 One Click Guided Automatic RECIST Lesion Measurement and Segmentation on CT Scans. *RSNA Annual Meeting*. Virtual, 2020-12-01.
- 6 Making RECIST Measurements Easy: A Semi-Automated Deep Learning System With Expert-Equivalent Accuracy and Better Consistency. *RSNA Annual Meeting*. Virtual, 2020-12-03.
- 5 Accurate Liver and Tumor Segmentation on CT Scans Using an Edge-enhanced Network. *RSNA Annual Meeting*. Virtual, 2020-12-04.
- 4 One Click Lesion RECIST Measurement and Segmentation on CT Scans. *International Conference on Medical Image Computing & Computer Assisted Intervention*. Virtual, 2020-10-06.
- 3 E<sup>2</sup>Net: An Edge Enhanced Network for Accurate Liver and Tumor Segmentation on CT Scans. *International Conference on Medical Image Computing & Computer Assisted Intervention*. Virtual, 2020-10-06.
- 2 Deep Learning in Radiology: Applications in Lesion and Organ Segmentation. *NIH AI Workshop on Image Segmentation*. Bethesda, USA, 2019-02-14.
- 1 Text-independent Writer Identification via CNN Features and Joint Bayesian. *International Conference on Frontiers in Handwriting Recognition*. Shenzhen, China, 2016-10-24.