CHI IAN TANG

Research Scientist in AI & Wearable Computing

Nokia Bell Labs, Cambridge, UK • iantangcam@gmail.com • https://iantang.co/

CAREER SUMMARY

Research Scientist with a PhD in Computer Science from the University of Cambridge, specializing in AI/ML for wearables and smart devices. Expertise in developing efficient and novel AI algorithms for resource-constrained applications, including recognizing human activities from wearable sensors (IMUs). Proven track record in industrial and academic research, published in top-tier venues (ICML, WACV, UbiComp, ACM IMWUT, ICASSP). Leads multi-modal AI research, integrating inertial sensors with text and vision data for wearable-based health analysis and other applications utilizing on-device LLMs. Proficient in PyTorch, with a demonstrated ability to lead advanced research projects and collaborate across interdisciplinary teams to develop scalable AI solutions.

EXPERIENCE

Research Scientist

Nokia Bell Labs, Cambridge — 09/2023 – Present

- Developing efficient, collaborative, and adaptive AI platforms for the future of personal devices
- Leading research efforts on cross-modality analysis (inertial sensors, text, vision) and contrastive-based label-efficient modeling (CLIP-based) for human motion and behavior
- Initiating research on multi-modal health analysis with on-device LLMs – integrating sensors with efficient on-device large language models for longitudinal health insights
- Supervising and leading PhD-level projects in wearable sensing and machine learning

Supervisor & Demonstrator

University of Cambridge — 11/2020 – Present

- Supervising and tutoring undergraduate students, covering Computation Theory, ML and Real-world Data

Research Intern

Nokia Bell Labs, Cambridge — 06/2022 – 09/2022

- Developed a label-efficient continual learning framework that can mitigate catastrophic forgetting and balance different learning objectives with a carefully designed loss function (Published in WACV 2024)

Research Assistant

University of Hong Kong, Bioinformatics Algorithms and Core Technology Lab — 06/2018 – 07/2018

- Enhanced state-of-the-art performance in variant calling in single molecule sequencing from noisy data (3GS) by utilizing sequence models and novel model architectures (Published in Nat. Mach. Intell. 2020)
- Explored a wide range of model architectures and optimized data processing pipeline, while developing reusable utility functions for visualizing information captured by deep learning models

EDUCATION

PhD in Computers Science

Supervisor: Prof Cecilia Mascolo

University of Cambridge — 2019 – 2024

- Thesis: Self-supervised Learning for Data-efficient Human Activity Recognition
- Funded by Cambridge Trust, Doris Zimmern Charitable Foundation, and Nokia Corporation

MPhil in Advanced Computer Science — Distinction

University of Cambridge — 2018 – 2019

- Thesis: Collaborative Activity Recognition between Multiple Devices

BEng in Computer Science — First Class Honors

University of Hong Kong — 2014 – 2018

- Full Scholarship, Major GPA: 4.15/4.30 (Top Score), Overall GPA: 4.07

SELECTED PROJECTS

Longitudinal Health Insights using Wearable Devices and Efficient LLMs

Ongoing Research, to appear in TSALM Workshop @ NeurIPS 2024

- Research on efficient, privacy-preserving systems that provide users and healthcare professionals with longitudinal health insights, by leveraging user behavior patterns captured by embedded sensors on wearable devices, and the multi-modal reasoning capability of LLMs
- Developed a pre-training framework for IMU encoders leveraging ego-centric vision and text annotations, allowing further research in improving reasoning capabilities of LLMs with sensor time series data (to appear in NeurIPS 2024 TSALM workshop, full paper under review)

Scalable and Data-efficient Machine Learning for Human Activity Recognition

Appeared in ACM IMWUT 2021, 2022, ML4MH @ NeurIPS 2020, HCRL @ AAAI 2024

- Development of novel training algorithms that improve both data and label efficiency while ensuring robust recognition in mobile applications
- Proposed novel semi-supervised and self-supervised learning techniques for human activity recognition
  - Utilized synchronized data from different devices capturing the same activity as a natural source of transformation to replace hand-picked transformation for contrastive learning (ACM IMWUT 2022)
  - Proposed a novel architecture combining self-supervised training and self-training for HAR. Utilized abundant unlabeled data to complement scarce labeled data using self-training (ACM IMWUT 2021)
  - Adapted a contrastive learning technique, SimCLR, to mobile sensing. Observed that performance differs significantly with different transformation functions (NeurIPS 2020 ML4MH Workshop)

Overcoming Catastrophic Forgetting in Continual Learning

Appeared in WACV 2024, ICASSP 2022

- Development of approaches that balance stability and plasticity, ensuring models can generalize effectively across new tasks while retaining performance on past tasks
  - Proposal of a contrastive semi-supervised learning framework leveraging unlabeled data and occasional labeled data to balance different continual learning objectives (WACV 2024)
  - Adopted multitask learning during base model training to improve feature generalizability. Demonstrated that a more transferable base model is beneficial to continual learning (ICASSP 2022)

Federated Learning for Scalable Learning Algorithms

Appeared in ICML 2022

- Proposed an unsupervised federated learning architecture that orchestrates a distributed clustering task
- Demonstrated that the technique is scalable, efficient, and robust to statistical heterogeneity
- Proved that it guarantees good generalization performance under a linear probe

TECHNICAL SKILLS

ML frameworks	PyTorch, Tensorflow, LiteRT (TFLite), TensorFlow.js
Development	Python, Java (Android/Wear OS)

SELECTED PUBLICATIONS

1. Das, A. M., Tang, C. I., Kawsar, F., & Malekzadeh, M. (2024). PRIMUS: Pretraining IMU Encoders with Multimodal and Self-Supervisied Learning. To Appear In NeurIPS 2024 Workshop on Time Series in the Age of Large Models.
2. Tang, C. I., Qendro, L., Spathis, D., Kawsar, F., Mascolo, C., and Mathur, A. (2024). Kaizen: Practical self-supervised continual learning with continual fine-tuning. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pages 2841–2850.
3. Haresamudram, H., Tang, C. I., Suh, S., Lukowicz, P., and Ploetz, T. (2023). Solving the sensor-based activity recognition problem (soar): Self-supervised, multi-modal recognition of activities from wearable sensors. In Adjunct Proceedings of the 2023 ACM International Joint Conference on Pervasive and Ubiquitous Computing & the 2023 ACM International Symposium on Wearable Computing, pages 759–761.
4. Lubana, E. S., Tang, C. I., Kawsar, F., Dick, R. P., & Mathur, A. (2022). Orchestra: Unsupervised Federated Learning via Globally Consistent Clustering. In the Thirty-ninth International Conference on Machine Learning (ICML 2022), July 2022, Baltimore MD, USA.
5. Ma, D.*, Tang, C. I.*, & Mascolo, C. (2022). Improving Feature Generalizability with Multitask Learning in Class Incremental Learning. In ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 4173-4177). IEEE.
6. Jain, Y.*, Tang, C. I.*, Min, C., Kawsar, F., & Mathur, A. (2022). ColloSSL: Collaborative Self-Supervised Learning for Human Activity Recognition. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 6(1), 1-28.
7. Shah, K., Spathis, D., Tang, C. I., & Mascolo, C. (2021). Evaluating Contrastive Learning on Wearable Timeseries for Downstream Clinical Outcomes. In Machine Learning for Health (ML4H) 2021 Symposium, December 2021, Virtual.
8. Tang, C. I., Perez-Pozuelo, I., Spathis, D., Brage, S., Wareham, N., & Mascolo, C. (2021). SelfHAR: Improving Human Activity Recognition through Self-training with Unlabeled Data. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 5(1), 1-30.
9. Luo, R., Wong, C. L., Wong, Y. S., Tang, C. I., Liu, C. M., Leung, C. M., & Lam, T. W. (2020). Exploring the limit of using a deep neural network on pileup data for germline variant calling. Nature Machine Intelligence, 2(4), 220-227.

AWARDS

Doris Zimmern HKU Hughes Hall, Cambridge International & Nokia Corporation Scholarship	2019
Doris Zimmern HKU-Cambridge Hughes Hall Scholarships	2018
Dean's Honours List – University of Hong Kong	2015, 2016, 2018
Undergraduate Research Fellowship Programme – University of Hong Kong	2018
ACM-ICPC Asia Regional Contest Qingdao Site 2017 – Silver Medal	2017
Hong Kong IET Prize 2016	2016
ACM-HK Collegiate Programming Contest 2016 – Fourth Place	2016