Sihan Tan

WebPage • Google Scholar

Research Interest

My research interest lies in **Multimodal Understanding and Generation**, with a focus on **vision and language modeling** in real-world communication. I explore how modalities such as vision, speech, and text can be effectively aligned to support natural and accessible human-computer interaction.
As a representative challenge, I am currently working on **Sign Language Understanding**, where the input consists of fine-grained visual gestures and the output is semantic spoken text. This task not only requires temporal and cross-modal modeling but also reflects my broader goal: to develop **inclusive AI** that supports diverse human expression and enables more **accessible communication**.
Fields: Natural Language Processing, Computer Vision, Multimodal, Machine Learning
Topics: Sign Language Understanding, Multilingual Machine Translation (MT), Efficient MT Training

Professional Experience

University of Zurich — Zurich, Switzerland
Visiting Researcher — 08/2025 - Present
- Visit Language, Technology and Accessiblity Group under Prof. Sarah Ebling

Institute of Science Tokyo — Tokyo, Japan
Research Assistant — 09/2023 - Present
- Lead Sign Language Processing Team and deploy system for real-life sign language-based HRI
- Mentor master's, bachelor's, and exchange students

Research Student (formerly Tokyo Tech) — 10/2021 - 03/2022
- Non-degree program for Research
- worked on speech recognition using ESPnet

NHK Science and Technology Research Labortories — Tokyo, Japan
Visiting Researcher — 04/2024 - Present
- Work with NHK Sign Language Team
- Focus on multilingual translation, efficient training for low-resource sign language

Intern — 08/2023 - 03/2024
- Worked on Sign Language Translation under Taro Miyazaki
- Boosted Lab-level collaboration

Dalian University of Technology — Dalian, China
Research Assistant — 12/2019 - 06/2022
- Worked on WiFi signal-based Human Action Recognition via cross-modal

EDUCATION

Institute of Science Tokyo (Formerly Tokyo Tech) — Tokyo, Japan
Ph.D. in Systems and Control Engineering — 04/2024 - EST. 03/2027
Advisor: Prof. Kazuhiro Nakadai

M.E. in Systems and Control Engineering, *Best Graduate* — 03/2022 - 03/2024
Advisor: Prof. Kazuhiro Nakadai, Prof. Katsutoshi Itoyama

Dalian University of Technology — Dalian, China
B.E. in Digital Media Technolgy, *Outstanding Bachelor's Graduate* — 09/2017 - 07/2021
Advisor: Xin Fan

Publication

Journal

[J3] **A Review of Deep Learning-based Approaches to Sign Language Processing**

Sihan Tan, Nabeela Khan, Zhaoyi An, Yoshitaka Ando, Rei Kawakami, Kazuhiro Nakadai
Advanced Robotics, 2024, December, 1–19.

[J2] **Advancing Human-Computer Interaction: End-to-End Sign Language Translation**

Sihan Tan, Katsutoshi Itoyama, Kazuhiro Nakadai
The Transactions of Human Interface Society, 2024 Volume 26 Issue 4 Pages 391-398.

[J1] **Motion Inbetweening Based on Body Parts Integration for Sign Language Generation**

Nabeela Khan, **Sihan Tan**, Katsutoshi Itoyama, Kazuhiro Nakadai
The Transactions of Human Interface Society, 2024 Volume 26 Issue 4 Pages 431-442.

Conference

[C5] **SignFlow: End-to-End Sign Language Generation for One-to-Many Modeling using Conditional Flow Matching**

Nabeela Khan, Bowen Wu, **Sihan Tan**, Carlos Toshinori Ishi, Kazuhiro Nakadai
In *Proceedings of the 27th ACM International Conference on Multimodal Interaction* (ICMI), 2025.

[C4] **Towards Online Sign Language Expression for Real-Time Human-Robot Interaction**

Nabeela Khan, **Sihan Tan**, Kazuhiro Nakadai
In *Proceedings of the 34th IEEE International Conference on Robot and Human Interactive Communication* (RO-MAN), 2025.

[C3] **Multilingual Gloss-free Sign Language Translation: Towards Building a Sign Language Foundation Model**
Sihan Tan, Taro Miyazaki, Kazuhiro Nakadai.
In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics* (ACL), 2025.

[C2] **Improvement in Sign Language Translation Using Text CTC Alignment**
Sihan Tan, Taro Miyazaki, Nabeela Khan, Kazuhiro Nakadai.
In *Proceedings of the 31st International Conference on Computational Linguistics* (COLING), 2025.

[C1] **Improving Sign Language Understanding Introducing Label Smoothing**
Sihan Tan, Nabeela Khan, Katsutoshi Itoyama, Kazuhiro Nakadai
In *Proceedings of the 32nd IEEE International Conference on Robot and Human Interactive Communication* (RO-MAN), 2023.

Workshop

[W2] **SEDA: Simple and Effective Data Augmentation for Sign Language Understanding**

Sihan Tan, Taro Miyazaki, Katsutoshi Itoyama, Kazuhiro Nakadai.
In *Proceedings of the LREC-COLING 11th Workshop on the Representation and Processing of Sign Languages* (sign-lang), 2024.

[W1] **Sign Language Translation with Gloss Pair Encoding**
Taro Miyazaki, **Sihan Tan**, Tsubasa Uchida, and Hiroyuki Kaneko
In *Proceedings of the LREC-COLING 11th Workshop on the Representation and Processing of Sign Languages* (sign-lang), 2024.

Grants

Heyning-Roelli Foundation Scholarship 08/2025 - 01/2026
Research Fund, stipend
Publisher: Heyning-Roelli Foundation

Global Off-Campus Project 08/2025 - 01/2026
Research Fund, stipend
Publisher: Academy of Super Smart Society, Science Tokyo

Academy of Super Smart Society Scholarship 06/2023 - present
Stipend
Publisher: Academy of Super Smart Society, Science Tokyo

Tsubame Special Scholarship (top 20%) 04/2024 -present
Stipend
Publisher: Institute of Science Tokyo

Awards

Miura Award, The Japan Society of Mechanical Engineers. 03/2024
Given to the most outstanding graduate student in the department.

Department Prize for Outstanding Paper Presentation, Tokyo Institute of Technology. 03/2024
Awarded for master's thesis presentation.

Misc.

Professional Service

- **Peer Review:** ROMAN (2023, 2025)
- **Journal Reviewer**: IEEE Access

Mentoring

- **Master's student:** Continuous Sign Language Recognition
- **Bacholar's student:** Customized Sign Language Translation
- **Exchange student:** Online Sign Language Processing System

Skills

Languages: Mandarin (Native), English (Fluent), Japanese (Fluent), German (Basic)
Coding: Python, PyTorch, TensorFlow, Linux, Matlab, C, C++, SQL, LaTeX, . . .