

Wanjiang Weng

Email: wjweng@seu.edu.cn / Page: [Github](#) / Phone: +86-15855089405

Address: 1 Jiulonghu, Jiangning, Nanjing, Jiangsu, China

EDUCATION BACKGROUND

Southeast University

Aug.2023-Jun.2026

- Major: Computer Technology
- Degree: Master of Engineering
- IETLS: **Overall-6.5**, L-6.5, R-7.5, W-6.5, S-5,
- GPA: 3.17/5.00; 3.15/4.00 Avg. Score: 81.69/100.00

Anhui University

Aug.2019-Jun.2023

- Major: Data Science and Big Data Technology
- Degree: Bachelor of Engineering
- GPA: 3.13/5.00; 3.36/4.00 Avg. Score: 83.38/100.00

RESEARCH INTEREST

- Human-Centric Perception Sep.2023-Present
- Skeleton-based Human Action Understanding Sep.2023-Present
- AIGC Oct.2024-Present
- Text-to-Motion Generation Oct.2024-Present

Research Experience

The Research Project: Text-To-Motion Generation

Oct.2024-Present

- **Post-training for T2M Diffusion Models:** Propose EasyTune, a fine-tuning framework for diffusion models that decouples recursive dependencies and enables (1) dense and effective optimization, (2) memory-efficient training, and (3) fine-grained alignment. The paper was Accepted by ICLR26.
- **Alignment of T2M Diffusion models:** Propose a plug-and-play reward-guided alignment mechanism, ReAlign, to enhance semantic consistency and motion quality in text-to-motion generation. Dynamically guiding the diffusion process with a reward model effectively optimizes the results. The paper was Accepted by AAAI26.
- **Multilingual T2M Generation:** Proposed the first bilingual T2M dataset, BiHumanML3D, which addresses language imbalance through cross-lingual alignment and establishes a unified conditioning space, improving motion generation quality in zero-shot and code-switching settings. Paper submitted to ACM MM 26.

The Research Project: Skeleton-based Action Understanding

Sep.2023-Present

- **Test-Time Adaptation for ZSAR:** Propose OPA, a training-free online prototype adaptation framework designed to bridge the inherent modality gap between textual prototypes and visual skeleton features. OPA enables dynamic and robust estimation of the true visual centers for unseen categories. Paper submitted to ACM MM 26.
- **Skeleton Representation Learning:** Propose a simple yet effective self-supervised skeleton dense representation learning method called USDRL, learns dense representation through multi-grained feature decorrelation. Design a dense skeleton encoder for dense prediction tasks. The paper was Accepted by AAAI25.
- Strong USDRL with Multi-view and multi-modal self-training and cover over eight action understanding tasks including action segmentation and early action prediction. The paper was Accepted by TPAMI25.

The Research Project: VLMs Inference Acceleration*May.2025-Jul.2025*

- **Speculative Sampling:** Proposed an autoregressive image-generation acceleration method, PEANUT, which dynamically adjusts draft-tree shapes based on pixel-block depth and width to mitigate draft-tree imbalance and improve sampling efficiency. Under Review.

PUBLICATION (* denotes Co-first Author, † denotes Corresponding Author)

- Tan X*, **Weng W***, Wang H†, et al. "EasyTune: Efficient Step-Aware Fine-Tuning for Diffusion-Based Motion Generation." ([ICLR26](#), Accepted).
- **Weng W***, Tan X*, Wang H†, et al. "ReAlign: Text-to-Motion Generation via Step-Aware Reward-Guided Alignment." ([AAAI26](#), Accepted).
- Wang H†, **Weng W**, Geng X, et al. "Foundation Model for Skeleton-Based Human Action Understanding" ([TPAMI25](#), Accepted).
- **Weng W**, Wang H†, Wang J, et al. "USDRL: Unified Skeleton-based Dense Representation Learning with multi-grained feature decorrelation." ([AAAI25](#), Accepted).
- **Weng W**, Wu Y, Wang H†, et al. "Bridging the Modality Gap for Zero-Shot Action Recognition via Online Prototype Adaptation." (Under Review).
- **Weng W***, Tan X*, Wang H†, et al. "Bilingual Text-to-Motion Generation: A New Benchmark and Baselines." ([Under Review](#)).
- Tan X*, **Weng W***, Wang H†, et al. "MotionRFT: Unified Reinforcement Fine-Tuning for Text-to-Motion Generation." ([Under Review](#)).
- Lei H*, **Weng W***, Wang H†, et al. "PEANUT: Fast Inference of Visual Autoregressive Model with Adjacency-Adaptive Dynamic Draft Trees." (Under Review).

INTERNSHIP EXPERIENCES**Hi-Think** —*Machine Learning Engineer**Jun.2025-Sep.2025*

- Responsible for optimizing speculative sampling algorithms (Medusa, Eagle, RAG-based), achieving 4× faster inference and 5× higher throughput without performance loss. Built an extensible Axolotl-based framework to support rapid adaptation and iteration across business models.

Sea-AI Lab —*Machine Learning Researcher**Feb.2026-Present*

- Unified Generation & Understanding: Research focuses on Unified Multimodal Models, including image generation, thinking with image. Experienced in training foundational models on large-scale GPU clusters.

EXTRACURRICULAR ACTIVITIES**ACM Lab of Anhui University** —*Fervent Member**Aug.2020-Jun.2022*

- Participated in multifarious competitions pertaining to programming and algorithms
- Obtained a number of prizes on behalf of Anhui University composed of The 2021 ICPC Asia Shanghai Regional Contest(Nomination), the 2020 Group Programming Ladder Tournament(the Outstanding Award of Higher Education Universities in Anhui Province), and so forth

AWARDS

National Scholarship for Postgraduates (Southeast University)	2025
Second Prize at the national level in National Post-Graduate Mathematical Contest in Modeling	2023
Bronze Medal of the 3 rd China Collegiate Algorithm Design & Programming Challenge Contest	2022
Bronze Medal in the 2021Collegiate Computer System & Programming Contest (East China Division)	2021
Honorable Mention of the 2021 ICPC Asia Shanghai Regional Contest and Jinan Regional Contest	2021