Jaewoo (Jeffrey) Heo

jeffheo@stanford.edu ▪ 650-788-7967 ▪ https://www.linkedin.com/in/jaewoo-jeffrey-h ▪ https://jeffheo.github.io/

EDUCATION

Stanford University

M.S in Computer Science (CS), Artificial Intelligence Track — Stanford, CA, **Expected June 2025**

B.S in Computer Science (CS) with Honors, Artificial Intelligence Track — Stanford, CA, **June 2024**

PROFESSIONAL EXPERIENCE

Medical AI and Computer Vision Lab (MARVL) — **Mar. 2023 – Present**

Researcher — Stanford, CA

- Led three first-author publications to CV/ML conferences and medical journals on 3D human motion reconstruction, video understanding, and medical/health AI
- Co-authored key research in leveraging VLMs for synthetic data generation, benchmarking VLMs on challenging medical/surgical tasks, and combining SAM2 with surgeon psychomotor metric algorithm for surgical tool tracking
- Spearheaded interdisciplinary medical AI research in surgical tool tracking (Wellcome LEAP:SAVE), and "Enhanced Trauma Video Review with Computer Vision" through collaborating with the clinical team

Stanford Vision and Learning Lab (SVL) — **Nov. 2023 – Present**

Researcher — Stanford, CA

- Developed a novel 3D HMR framework that leveraged NeRF rendering of the human body in "NeuHMR: Neural Rendering-Guided Human Motion Reconstruction"
- Leading a first-author project on monocular 3D HMR via fine-tuning the pre-trained Stable Diffusion U-Net

Ermon Group — **Mar. 2023 – Jan. 2024**

Student Researcher — Stanford, CA

- Led red-teaming project on the algorithmic generation of prompt injections as adversarial attacks on LLMs
- Generated effective prompt injection tokens that successfully hack open-source LLMs with at least 1.8% probability by using GPT-3 API, Llama-2, and Falcon-40B with a black-box algorithm

PUBLICATION

- **Heo, J.**, Hu, G., Wang, Z., & Yeung-Levy, S. (2025). DeforHMR: Vision Transformer with Deformable Cross-Attention for 3D Human Mesh Recovery. *International Conference on 3D Vision.*
  - Developed novel query-agnostic deformable cross-attention mechanism that allows the model to attend to relevant spatial features coherently, flexibly, and in a data-dependent manner
  - Achieved SOTA performance on 3D HMR benchmarks 3DPW and RICH by 7.8% and 10.3% improvement, respectively, from previous SOTA metrics
- **Heo, J.**, Wang, K., Liu, K., & Yeung-Levy, S. (2025). Motion Diffusion-Guided 3D Global HMR from a Dynamic Camera. *ArXiv.*
  - Developed a novel global 3D human motion reconstruction algorithm that leverages motion diffusion models and score distillation sampling to decouple and optimize both human and camera motion
  - Achieved SOTA performance on global 3D HMR benchmarks EMDB and Egobody by 15.9% and 19.7% improvement, respectively from 2nd best algorithm
- Bravo-Sanchez, L., **Heo, J.**, Weng, Z., Wang, K., & Yeung-Levy, S. (2025). Ask, Pose, Unite: Scaling Data Acquisition for Close Interactions with Vision Language Models. *ArXiv.*
  - Designed a novel data generation method for close interactions that leverages automatic annotations from VLMs to scale data acquisition, producing pseudo-ground truth meshes from in-the-wild images
- Xiang, T., Wang, K., **Heo, J.**, Adeli, E., Yeung-Levy, S., Delp, S., & Fei-Fei, L. (2025). NeuHMR: Neural Rendering-Guided Human Motion Reconstruction. *International Conference on 3D Vision.*
  - Developed an optimization-based 3D HMR framework based on neural rendering, rethinking the dependency on the 2D key-point fitting paradigm
- Villarreal, J., **Heo, J. (co. first-author)**, Wang, X., Bain, A., Succar, B., Yao, D., Jopling, J., Yeung-Levy, S., Dumas, R. (2025) Enhanced Trauma Video Review with Computer Vision – Trauma Resuscitation Phase Segmentation and Procedure Detection. *Under Review.*

Jaewoo (Jeffrey) Heo

jeffheo@stanford.edu ▪ 650-788-7967 ▪ https://www.linkedin.com/in/jaewoo-jeffrey-h ▪ https://jeffheo.github.io/

    - Developed a novel CV pipeline involving video feature extraction followed by downstream models that temporally segment resuscitation phases and detect key medical procedures
- Rau, A., Endo, M., Aklilu, J., **Heo, J.**, Saab, K., Paderno, A., Jopling, J., Holsinger, F., Yeung-Levy, S. (2025) Systematic Evaluation of Large Vision-Language Models for Surgical Artificial Intelligence. *Under Review.*
    - Evaluated 11 state-of-the-art VLMs across 17 surgical AI tasks using 13 datasets, demonstrating VLMs' superior generalizability compared to supervised models when deployed outside their training environments

PROJECT EXPERIENCE

Wellcome LEAP:SAVE (SOAR Computer Vision Team, MARVL) **Jun. 2024 – Present**

- Developed a novel surgical tool-tracking algorithm and automatic surgical performance metric calculation algorithm to compute surgeons' performance metrics on laparoscopic cholecystectomy videos
- Conducting correlation study on computed metrics and surgeons' corresponding meta-competency annotations

Stable-Diffusion Fine-tuning with Foundation Model Feature Conditioning for 3D HMR (SVL) **Aug. 2024 – Present**

- Fine-tuning a pre-trained Stable Diffusion U-Net with DINOv2 and ViTPose features as conditioning for 3D HMR

TECHNICAL SKILLS & INTERESTS

Skills: 3D human mesh recovery, 4D motion capture, video understanding, VLM, LLM, conditional diffusion model, score distillation sampling, medical image segmentation, temporal action segmentation, action detection, surgical tool segmentation and tracking, Stable Diffusion fine-tuning, synthetic data generation

Tech Stack: Python (PyTorch, PyTorch Lightning, CUDA, Keras, NumPy, Pandas, HF Transformers), C/C++, Java, MATLAB, Julia, R, Unix

Interests: UEFA Champions League, Tennis, Guitar, Music Production, The Beatles, Billy Joel, The Office, One Piece