DAYE LEE

1, Gwanak-ro, Gwanak-gu, Seoul, Republic of Korea
(82)-10-5542-0631 | dayelee313@gmail.com

EDUCATION

Seoul National University Seoul, Korea
Master of Data Science *Mar 2022 - Expected in Aug 2025*

- Multi-modal Deep learning Theories and Application, Sep 23 - Dec 23
- Natural Language Processing, Conversational AI, Dialogue Systems, Sep 23 - Dec 23

University of Toronto Toronto, Canada
Study Abroad, Centre for Analytics and Artificial Intelligence Engineering *Jan 2024 - Jun 2024*

- *Sponsored by Institute for Information & Communication Technology Planning & Evaluation (IITP), Korean Government Institute*
- Mobile Robotics and Perception, Jan 24 - Apr 24
- Introduction to Deep Learning, Jan 24 - Apr 24
- Data Science Methods and Statistical Leearning, Jan 24 - Apr 24

Seoul National University Seoul, Korea
B.S in Applied Biology and Chemistry *Feb 2022*

- GPA: 3.72/4.3

Cadillac High School Michigan, U.S.
Study Abroad *2013 - 2014*

TECHNICAL SKILLS

Languages (Advanced): Python, C/C++
Languages (Intermediate): SQL(PostgreSQL), MATLAB
Developer Tools: Git, VS Code, Wandb
Libraries & Frameworks: Pytorch, Pytorch3D, Accelerate, ROMP, VIBE, Linux, ROS
Software tool: Blender, RViz

PROFESSIONAL PUBLICATIONS

Daye Lee, Minseok Kong, Jiho Park, Nikolaos Kourtzanidis et al. 2025. *Simulating Mobile Robot Vision: An Analysis of RGB-D versus RGB-Based Accuracy and CPU Optimization*. ICAIIC

Daye Lee, Taesup Kim. 2025. *HuMoGen-X: User-Guided Editing, Genre-Aware 3D Motion Generation via Music-Conditioned Diffusion.* Master's Thesis, Seoul National University. (In preparation for CVPR 2026 submission)

EXPERIENCE

Cyberworks Robotics, Autonomous self-driving company Toronto, Canada
Project member of three students *Jan 2024 – Jun 2024*

- Processed live RGB camera input by consecutively applying YOLOv8 and a monocular depth estimation model, MobileNet2, integrating their outputs, including 2D bounding box coordinates and estimated depth information.

- Visualized results in real-time using RVIZ with 3D bounding box markers (Cube and Arrow) and Bird's-Eye View (BEV) visualizations to verify the accuracy of estimated depth.
- Achieved a 7% reduction in CPU usage for the RGBD model and a 5% reduction for the RGB model compared to non-optimized versions as a team.
- Developed an efficient real-time human detection system for low-computing-resource robots equipped with CPU-only capabilities, such as indoor wheelchairs or large cleaning robots.

PROJECTS

3D Motion Generation | *Diffusion, Blender, ROMP, JukeBox* Sep 2024 – Aug 2025

- Proposed a controllable diffusion framework combining **music, genre, and keyframe** conditions for 3D dance motion synthesis.
- Introduced **Chained Classifier-Free Guidance (CFG)**, improving structural fidelity (MPJPE_k: 7.40 vs. 53.84 of POPDG, ↓ 86%) and genre expressivity (F1: 0.176 vs. 0.091 GT) while maintaining high Recall@0.3cm (0.4365)
- Constructed the **CrossGenreDanceSet** in SMPL (A Skinned Multi-Person Linear Model) format using ROMP, a 12-genre dataset with genre-mismatched pairs, demonstrating superior motion-music alignment (Beat Align: 0.3029 vs. 0.2499 in POPDG, +21%) and reduced motion jitter to 2.93 (vs. 4.82 of POPDG, ↓ 39%) and achieved higher motion diversity ($Dist_g$ = 7.58 vs. 6.27 EDGE).
- Implemented **Gaussian-masked keyframe conditioning**, reducing transition jitter by 39% compared to baselines and enhancing smoothness.
- Introduced several metrics for structural fidelity (MPJPE, $MPJPE_k$, $MPJPE_g$, Recall@0.3)
- Developed a dance genre recognition classifier trained on **GenreDanceSet** (120 min curated data), improving genre expressivity over ground truth (Top-3 accuracy: 0.528 vs. 0.307, F1: 0.176 vs. 0.091).

Emotion-Specialized Text-to-Video Retrieval | *Python, Pytorch* Jan 2024 – May 2024

- Developed a text-to-video retrieval system that specializes in extracting videos based on emotion-rich text queries
- Modified code to align with available distributed training library - accelerate, conducted model training experiments, and shared fine-tuned model checkpoints with the team
- Implemented evaluation metrics such as recall based on cosine similarities matrix to measure the similarity between the text query and the video embeddings in the dataset, retrieving the top-n videos in descending order of relevance
- Discussed the t-SNE result of embedding space and identified changes in the embeddings associated with similar emotions to be mapped closer together in the space
- Found a decline in overall quantitative performance when contrasting baseline with our methods, concluding that less training dataset led to the performance degradation
- Implemented demonstration code in Jupyter Notebook to showcase the project

Personalization of Music Conditioned Dance generation | *Python, Pytorch* Sep 2023 – Dec 2023

- Deployed few-shot learning methods such as Dreambooth, Inversion and etc, and fine-tuned the EDGE model in three approaches with three different few-shot dataset, a Diffusion-based baseline model where the model learns the correlation between music and SMPL pose data through cross-attention
- Generated few-shot dataset by crawling YouTube and processed the data with ROMP to extract SMPL information
- Resulted in decreasing FID_k by 81, FID_m by 2.432 and PFC by 0.98 and enhancing Beat Alignment by 0.074 and outperforming in human evaluation in all three approaches
- Demonstrated the qualitative motion results by rendering them using Blender

Dialog Inpainting for Legal Dialogue Systems | *Python* Sep 2023 – Dec 2023

- Utilized the dialog inpainting approach (Dai et al., 2022) as a data augmentation technique to construct a dataset suitable for developing a legal consultation chatbot.
- Identified three key characteristics of legal documents: hierarchical structure, inter-referencing structure, and similar-context structure.
- Created a question-answer dialogue dataset using the ChatGPT 3.5 model, generating 3,048, 3,408, and 474 conversations for the hierarchical, referencing, and similarity-based approaches, respectively.
- Evaluated the fine-tuned Llama2-7B chat model against a baseline model based on accuracy, informativeness, well-formedness, and overall quality across 10 conversations for the question classification (QC) versus question answering (QA) task and 15 conversations for legal domain-specific approaches, involving evaluation by all five team members.

VOLUNTEER AND TEACHING EXPERIENCE

TED Translators Sep 2024 – Present

- Translated Two TED Talks from English to Korean mainly of TED ED and TED original content

Introduction to Computing Lecture Tutor Mar 2025 – Aug 2025

- Tutored undergraduate students for two semesters in the course **Introduction to Computing: Foundations and Basics**.
- Recognized with the Outstanding Tutor Award for excellence in teaching support.

CERTIFICATES

PyTorch for Deep Learning with Python bootcamp Oct 2023

- NumPy, Pandas, Machine Learning, Model Evaluation, Tensors, Neural Network, ANN, CNN, RNN