

HAH MIN LEW

github.com/hahminlew | hahminlew.github.io | hahmin.lew@gmail.com | [LinkedIn](#) | [Google Scholar](#) | [+8210.6876.3175](#)

CORE VALUES

Data-centric approaches leading to ML systems can solve valuable real-world problems. I value building efficient systems with open-sourcing through agile trials and errors based on a clarified problem. Proactively growing, challenging, and sharing are my core values and attitude toward my life.

EDUCATION

Mar. 2019 - Aug. 2021 M.S. in Electrical Engineering & Computer Science at **DGIST** (GPA: 4.06/4.3)
 Mar. 2014 - Feb. 2019 Bachelor of Engineering at **DGIST** (Best Project Award)

EXPERIENCE

Klleon, AI Researcher Aug. 2022 - present

- Building a data construction pipeline and generative model frameworks for a virtual human dialogue system.

DGIST, Graduate Researcher, Advisor: Prof. Jae Youn Hwang Mar. 2019 - Aug. 2022

- Multimodal Biomedical Imaging and System Lab (MBIS Lab).
- [6 SCIE publications](#), 7 international conferences, 9 projects, 6 patents, and [2 awards](#).
- ML-based anomaly labeling and generative model design for various biomedical applications under multi-modal, class-imbalance, and multi-task problems.
- Frequency-domain translation for substantial and informative data creation through generative models.
- Developed a novel 1-D time-series signal processing algorithm for the biomedical monitoring system.
- Collaborative research experiences with medical doctors from SNUH, SNUHD, Yonsei Severance, etc.

LANTERN, Co-founder Nov. 2016 - July 2017

- Founded a data-driven personalized tutor matching service company. Co-working with **Class101**.

SELECTED PROJECTS

Building a virtual human dialogue system Aug. 2022 - present

- Developed conditional generative model frameworks to create photo-realistic virtual humans.
- Constructed a multimodal data collection and preprocessing pipeline, the state-of-the-art ML model, training and evaluation frameworks, and an inference pipeline.
- Improved service latency from 10 seconds to 16.5 milliseconds per frame and generative performances.
- Used skills: Python, PyTorch, Docker, Git

Finetuning a large text-to-image model with a custom-built dataset maker Oct. 2023 - Nov. 2023

- Full open source contributions of fashion dataset creation with an ML-based captioning module, finetuning demo codes, and inferable text-to-image models.
- Used skills: Python, PyTorch, Git | Repositories: [\[Github\]](#), [\[Model\]](#), [\[Dataset\]](#)

Smartphone-based image classification for detecting early dental caries Apr. 2020. - Feb. 2022

- Achieved 95.6% accuracy, 0.952 recall and 0.953 precision scores.
- Used skills: Python, TensorFlow

Multimodal data analysis for the mobile diagnosis of otitis media Feb. 2020. - Jan. 2022

- Constructed multimodal human clinical data into trainable matrices (up to 4.98 billion pixels).
- Enhanced diagnostic accuracy by a multi-layer perceptron (80%) exceeds that of expert clinicians (73%).
- Used skills: Python, TensorFlow

Multimodal data registration and analysis for tumor/cancer detection Mar. 2019. - Oct. 2020

- Pixel-wise spectral image classification for tumor characterization (sensitivity: 0.86, specificity: 0.85).
- Used skills: Python, MATLAB

SELECTED PUBLICATIONS

Hah Min Lew*, S. Yoo*, H. Kang*, et al., “Towards High-Fidelity Head Swapping with Chroma Keying”, Under Review in **CVPR 2024**. [\[Project Page\]](#)

- Design a novel foreground-prediction and -aware transformer and augmentation method to improve head swapping performances under a self-supervised training.

Hah Min Lew*, J. S. kim*, et al., “Deep Learning-based Synthetic High-Resolution In-Depth Imaging Using an Attachable Dual-element Endoscopic Ultrasound Probe”, Arxiv Preprint 2023. [\[Paper\]](#)

- Data-centric and fine-tuning approach for high-resolution medical image generation using generative models.

K. Lee, **Hah Min Lew**, et al., “CSS-Net: Classification and Substitution for Segmentation of Rotator Cuff Tear”, In **ACCV 2022**. [\[Paper\]](#)

- Developed a multi-task network for detection of class-imbalanced regions.
- Employed DFT-based frequency translation for substitution of positive and negative regions.

T. C. Cavalcanti, **Hah Min Lew**, et al., “Intelligent Smartphone-based Multimode Imaging Otoscope for the Mobile Diagnosis of Otitis Media”, Biomedical Optics Express (IF: 3.562, [Spotlight on Optics](#)), 2021. [\[Paper\]](#)

- Image classification via ML algorithms (*Multi-layer perceptron, Random forest, Logistic regression, Decision trees, Naïve Bayes*) for multimodal human data from clinical trials.
- Quantitative analysis using standard metrics.

Hah Min Lew, et al., “Ultrasonic Blood Flowmeter with a Novel Xero Algorithm for a Mechanical Circulatory Support System”, Ultrasonics (IF: 4.062), 2021. [\[Paper\]](#)

- Developed a novel signal processing algorithm that complements for conventional flow monitoring methods.

J. Kim, **Hah Min Lew**, et al., “Forward-looking Multimodal Endoscopic System based on Optical Multispectral and High-frequency Ultrasound Imaging Techniques for Tumor Detection”, **IEEE TMI (IF: 11.037)**, 2020. [\[Paper\]](#)

- Multimodal data collection of human colon tissues ex vivo.
- Classification of tumor from multispectral data through the spectral angle mapper algorithm.
- Proposed the multimodal tumor characterization system using both depth-wise and surface-wise data.

AWARDS

Outstanding Poster Award Aug. 2021

- 2021 Student Conference, DGIST







Outstanding Paper Award May. 2021

- 2021 Spring Conference, The Korean Society of Medical & Biological Engineering (KOSOMBE)

Best Project Award Mar. 2017

- 2016 Undergraduate Group Research Project (UGRP) Program, DGIST

SKILLS

Programming  Python,  PyTorch,  TensorFlow,  MATLAB |  Docker,  Git
Languages Korean (native), English (professional working proficiency)