

Hah Min Lew

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

CORE VALUES

A data-centric approach leading to ML systems can solve valuable real-world problems. In order to create a working business with a clarified problem, I value building efficient ML systems with open-sourcing through agile trials and errors. 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 dataset construction pipeline and a generative model framework for virtual human creation.
- Selected Tasks: Text-to-Image Generation, Finetuning, Vector Quantization.

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

- 6 SCIE publications, 7 international conferences, 9 projects, 6 patents, and 2 awards.
- ML-based anomaly labeling for various biomedical applications under multimodal, class-imbalance, and multitask 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, SNUDH, Yonsei Severance, etc.

LANTERN, Co-founder Nov. 2016 - July 2017

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

SELECTED PROJECTS

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

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

Building a state-of-the-art head swapping pipeline Aug. 2022 - June 2023

- Full cycle construction of from data collection and preprocessing automation pipeline, ML model design, training and evaluation, result serving and improvement.
- Increasing generative performance with 2.2 times faster FPS and 64% fewer parameters.
- Used skills: Python, PyTorch, Git

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

- Collected and constructed multimode data for detecting early dental caries.
- 10-fold cross-validation of classification performance using standard metrics.
- Used skills: Python, TensorFlow

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

- Constructed 4-D multimode human clinical data into trainable 2-D matrices (up to 4.98 billion pixels).
- Outperformed diagnosis accuracy of expert clinicians (73%) using a multilayer perceptron algorithm (80%).
- Used skills: Python, TensorFlow

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

- Aligned depth-wise data to corresponding surface-wise data through vector calculations.
- 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 multitask 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 (*Multilayer perceptron, Random forest, Logistic regression, Decision trees, Naïve Bayes*) for multimodal human data from clinical trials 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. [\[Paper\]](#)

- Data collection from proposed multimodal imaging of human colon tissues including tumors ex vivo.
- Conducted quantitative analysis of multimode data for tumor characterization.







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)