

# HAH MIN LEW

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

## CORE VALUES

Data-centric approaches leading to ML systems can solve valuable real-world problems. I value building efficient systems 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 (EECS) 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 efficient data preprocessing pipelines and generative models 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 detection and generative model design for biomedical applications under multimodal, class-imbalance, and multi-task problems.
- Development of application-specific systems integrating hardware and software.
- Advanced signal processing and computer vision skills.

**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

- Built a multimodal data preprocessing pipeline and the state-of-the-art ML model, integrated Distributed Data Parallel for training, devised in-house training and evaluation frameworks, and established an efficient inference pipeline.
- Improved inference latency from 10 seconds to 47 milliseconds per frame and generative performances.
- Sped up in-house data preprocessing pipeline from 3 FPS to 36 FPS.
- 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 product dataset creation with an ML-based captioning module, finetuning demo codes, and inferable text-to-image models.
- Used skills: Python, PyTorch, Git | Repositories: [\[Github\]](#), [\[Model\]](#), [\[Dataset\]](#)

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

- Constructed clinical data into trainable matrices (4.98 billion pixels) from a proposed mobile otoscope.
- Enhanced diagnostic accuracy by a multi-layer perceptron (80%) exceeds that of expert clinicians (73%).
- Used skills: Python, Scikit

**Development of a real-time biomedical monitoring system** Mar. 2019. - Mar. 2021

- Developed a cost-effective 1-D time-series signal processing algorithm of time complexity  $O(N \log N)$ .
- Used skills: MATLAB, LabView, VHDL

## SELECTED PUBLICATIONS

---

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

- Designed 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 finetuning 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.
- Proof reliability of generated data through t-SNE clustering.

M. H. Lee, **Hah Min Lew**, et al., “Deep learning-based framework for fast and accurate acoustic hologram generation”, IEEE TUFFC (IF: 3.267, [Frontal Cover Paper](#)), 2022. [\[Paper\]](#)

- Designed autoencoder architecture for unsupervised learning, loss functions for both accuracy and energy efficiency, and physical constraints layers for acoustic hologram generation in the real world.

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

- Image classification via fundamental 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 biomedical monitoring system integrating hardware configurations and a proposed algorithm.
- Experience FPGA programming to generate bipolar pulses with a clock frequency of 100 MHz and time intervals of 20  $\mu$ s to calculate real-time flow rates.

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. [\[Paper\]](#)

- Classification of tumor from multispectral data through the spectral angle mapper algorithm.
- Proposed a 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

---

Skills  Python,  MATLAB, Bash |  PyTorch,  TensorFlow, Scikit, OpenCV, Pandas |  Docker,  Git  
Languages Korean (native), English (professional working proficiency)