

# Haoyu Dong

haoyu.dong151@duke.edu | +1 858 500-2825 | Homepage | LinkedIn | Google Scholar

## Research Profile

---

Ph.D. candidate in Electrical & Computer Engineering at Duke University, specializing in foundation models, LLM agents, multimodal learning, and data-centric AI. Industry research experience at Meta and Siemens Healthineers, developing algorithm-evolution methods for LLM agents and in-context learning systems for 3D medical image segmentation. Author of 15+ peer-reviewed publications in medical AI, computer vision, and multimodal learning.

## Industry Experience

---

### Meta

May 2026 – Present

*Software Engineer Intern*

*Menlo Park, CA*

- Improved Meta-Harness’s algorithm-evolution pipeline by making agent trajectories evolvable and dynamically adapting the training set to strengthen the optimization signal. Increased performance from **49% to 62%** on math-solving tasks (**+13 percentage points**) and from **68% to 71.3%** on Terminal-Bench 2 (**+3.3 percentage points**).

### Siemens Healthineers

Jun. 2025 – Aug. 2025

*AI Research Intern*

*Princeton, NJ*

- Developed a training-free in-context segmentation framework with fixed support-query prompting; built preprocessing and evaluation infrastructure spanning **14K MRI/CT volumes**. Co-inventor on the resulting **patent-pending** technology assigned to Siemens Medical Solutions USA, Inc.
- With only five supports per task and no retraining, achieved **52.0% Dice** versus **60.0%** for a supervised nnU-Net on an internal test set, and **54.9%** versus **56.9%** on the AMOS CT validation set.

### Gradient Health

May 2020 – Sep. 2020

*Fellowship Recipient*

*Durham, NC*

- Built a private multimodal dataset of diagnostic images and paired radiology reports from **2.6M patients**; developed a language-informed BYOL pretraining pipeline that improved accuracy by **3 percentage points** over a ResNet baseline.

## Selected Research Experience

---

### Mazurowski’s Lab, Duke University

Jul. 2022 – Present

*Graduate Research Associate*

*Durham, NC*

- Pretrained **MRI-CORE** on **7M MRI slices** using modern self-supervised objectives, improving few-shot downstream segmentation by **5 Dice points** over SAM; currently developing a 3D multimodal breast MRI foundation model that aligns volumetric imaging with radiology reports.
- Led a broad evaluation of promptable segmentation across **19 datasets, 5 modalities, and 18 body locations**, establishing data-efficient interactive workflows that transfer across scanners and sites (*Medical Image Analysis*, 2023).
- Developed **InTEnt**, a single-image test-time adaptation method using integrated entropy weighting to adjust batch-normalization statistics, improving out-of-domain segmentation by **2.9 Dice points** over the strongest comparator (CVPR 2024 Workshop Oral).
- Co-developed foundation-model adaptation guidelines spanning prompt design, parameter-efficient adaptation, and light fine-tuning, and contributed to universal anatomical segmentation models across locations and acquisition protocols.

### Mazurowski’s Lab, Duke University

Jun. 2021 – Jun. 2022

*Associate in Research*

*Durham, NC*

- Proposed **SWSSL** for anomaly detection in high-resolution breast tomosynthesis and chest X-ray, surpassing prior state of the art by **5.0 AUROC points** on tomosynthesis (*IEEE Transactions on Medical Imaging*, 2023).

### Sapiro’s Lab, Duke University

Sep. 2019 – May 2021

*Graduate Research Assistant*

*Durham, NC*

- Built a visual-semantic space for text-conditioned image retrieval using iterative point alignment and manifold-distance search, improving recall by **2 points** on FashionIQ, **5 points** on CSS, and **30 points** on a new dataset (CVPR 2021 Workshop).

## Selected Publications

---

- Haoyu Dong**, Yuwen Chen, Hanxue Gu, Nicholas Konz, Yaqian Chen, Qihang Li, and Maciej A. Mazurowski. “MRI-CORE: A Foundation Model for Magnetic Resonance Imaging.” *Preprint*, 2025.
- Haoyu Dong**, Han Liu, Riqiang Gao, Bogdan Georgescu, Guillaume Chabin, Jianing Wang, and Sasa Grbić. “Revisiting In-Context Learning for 3D CT Segmentation of Unseen Anatomies.” *Submitting*.

## Selected Publications (continued)

- Hanxue Gu\*, **Haoyu Dong\***, Pedro R. A. S. Bassi, Wenxuan Li, Alan Yuille, Zongwei Zhou, Kang Wang, and Yang Yang. “Reports Are Not Captions: Understanding and Designing Report Supervision for Label-Efficient 3D Lesion Segmentation.” *Submitting*.
  - Hanxue Gu\*, **Haoyu Dong\***, Jichen Yang, and Maciej A. Mazurowski. “How to build the best medical image segmentation algorithm using foundation models: a comprehensive empirical study with Segment Anything Model.” *Machine Learning for Biomedical Imaging*, 2025.
  - **Haoyu Dong**, Nicholas Konz, Hanxue Gu, and Maciej A. Mazurowski. “Medical image segmentation with InTEnt: Integrated entropy weighting for single image test-time adaptation.” *CVPR Workshops*, 2024.
  - Maciej A. Mazurowski, **Haoyu Dong**, Hanxue Gu, Jichen Yang, Nicholas Konz, and Yixin Zhang. “Segment anything model for medical image analysis: an experimental study.” *Medical Image Analysis*, 2023.
  - **Haoyu Dong**, Yifan Zhang, Hanxue Gu, Nicholas Konz, Yixin Zhang, and Maciej A. Mazurowski. “SWSSL: Sliding window-based self-supervised learning for anomaly detection in high-resolution images.” *IEEE Transactions on Medical Imaging*, 2023.
- \* Equal contribution. Full publication list available on Google Scholar.

## Education

|                                                                                                                  |                                         |
|------------------------------------------------------------------------------------------------------------------|-----------------------------------------|
| <b>Duke University</b><br><i>Ph.D., Electrical &amp; Computer Engineering</i>                                    | Expected Jan. 2027<br><i>Durham, NC</i> |
| <b>Duke University</b><br><i>M.S., Computer Science</i>                                                          | Jun. 2021<br><i>Durham, NC</i>          |
| <b>University of California, San Diego</b><br><i>B.S., Mathematics–Computer Science; B.S., Cognitive Science</i> | Jun. 2019<br><i>La Jolla, CA</i>        |

## Technical Skills

**Programming:** Python, C++  
**Machine Learning:** PyTorch, distributed training, reinforcement learning, multimodal modeling  
**Infrastructure:** Linux, Git, Docker, Slurm, cloud platforms

## Academic Service

**Conference Reviewer:** ICLR, NeurIPS, ICML, CVPR, ICCV, ECCV, MICCAI, MIDL  
**Journal Reviewer:** IEEE TMI (Distinguished Reviewer, Silver Level 2023–24), IEEE JBHI, IEEE TNNLS, IEEE TBME, Medical Image Analysis, npj Digital Medicine, npj Artificial Intelligence