Haekyu Park

CS PhD student at Georgia Tech

haekyu@gatech.edu Curriculum Vitae Google Scholar

I'm a Ph.D. student in Computer Science at Georgia Tech, working with Dr. Polo Chau. My research goal is to address fundamental challenges in understanding how AI models works and what they have learned: How do we scalably discover and summarize concepts that a model has learned? How do such concepts evolve as the model is trained? And how to identify and explain vulnerabilities that the model is susceptible to? My research focuses on three complementary thrusts:

- Scalable Visual Discovery to Interpret DNN Mechanism
- Insights to Protect and Troubleshoot Models
- Scalable Interpretation of Concept Evolution in Deep Learning Training

Specifically, I create novel tools that enable interactive scalable discovery of concepts, evolutions, and vulnerabilities in deep learning. My research is supported by JPMorgan AI PhD Fellowship. I have been fortunate to work with amazing researchers, engineers, and scientists at stripe, Microsoft, NVIDIA., and intel..

Education

Georgia Institute of Technology
Ph.D., Computer Science
Advisor: Dr. Polo Chau
Aug 2018 - Present

Seoul National University
B.S., Computer Science and Engineering
Graduated with honors (Cum Laude)
Mar 2012 - Aug 2017

Appointments

Machine Learning Engineering Intern Jun 2022 - Aug 2022
Stripe, Seattle, WA
Mentor: Revanth Rameshkumar
Designed and developed a deep neural network model for identifying fraud transactions, resulting in x8 faster training and (expected) $4-5M/year increase in the volume of detected fraud transactions for customers compared to the previous model.

Research Intern Jun 2021 - Aug 2021
Microsoft Research, Redmond, WA
Mentor: Gonzalo Ramos

AI Infrastructure Software Intern May 2020 - Jul 2020
NVIDIA, Santa Clara, CA
Mentor: Joe Eaton, Brad Rees, Bartley Richardson
Developed a visual graph analytics, allowing for interactively running multiple graph algorithms in real-time on large graphs. Leveraged GPU acceleration for both data analysis and rendering side

Data Science Intern May 2019 - Aug 2019

NVIDIA, Austin, TX
Mentor: Bartley Richardson, Brad Rees, Joe Eaton
Internship results are integrated into NVIDIA RAPIDS team's KDD 2019 NVIDIA RAPIDS tutorial

Graduate Research Assistant Aug 2018 - Present
Georgia Institute of Technology, Atlanta, GA

Undergraduate Research Assistant Jun 2016 - Aug 2017
Seoul National University, Seoul, Republic of Korea

Honors and Awards

Rising Stars in EECS 2022
Rising Stars in EECS, Hosted at the University of Texas at Austin.

J.P.Morgan PhD Fellowship 2021
For my PhD Work "Human-centered AI: Interactive Scalable Interfaces for Trustworthy and Safe AI"

"Thank a Teacher" Award 2019
Center of Teaching & Learning (CTL), Georgia Institute of Technology

Moon-Jung Chung Scholarship 2019
KOCSEA (The Korean Computer Scientists and Engineers Association in America)

National Scholarship for Science and Engineering 2015
National Scholarship for Science and Engineering

Grants and Funding

WiML Travel Funding 2019
$550 Travel Funding
Women in Machine Learning Workshop, co-located with NeurIPS

Amazon AWS Research Grant 2018
Funded $5,000 in AWS cloud credits
Co-PIs: Nilaksh Das, Scott Freitas, Duen Horng Chau

Publications

NeuroMapper: In-browser Visualizer for Neural Network Training
Zhiyan Zhou, Kevin Li, Haekyu Park, Megan Dass, Austin P. Wright, Nilaksh Das, Duen Horng Chau
IEEE Visualization Conference (VIS), 2022
🗎 Paper ▶ Demo

Explaining Website Reliability by Visualizing Hyperlink Connectivity
Seongmin Lee, Sadia Afroz, Haekyu Park, Zijie J. Wang, Omar Shaikh, Vibhor Sehgal, Ankit Peshin, Duen Horng Chau
IEEE Visualization Conference (VIS), 2022
🗎 Paper

DetectorDetective: Investigating the Effects of Adversarial Examples on Object Detectors
Sivapriya Vellaichamy, Matthew Hull, Zijie J. Wang, Nilaksh Das, Sheng-Yun Peng, Haekyu Park, Duen Horng Chau

Conference on Computer Vision and Pattern Recognition (CVPR), Demo, 2022
▶ Demo

MisVis: Explaining Web Misinformation Connections via Visual Summary
Seongmin Lee, Sadia Afroz, Haekyu Park, Zijie J. Wang, Omar Shaikh, Vibhor Sehgal, Ankit Peshin, Duen Horng Chau
CHI Conference on Human Factors in Computing Systems Extended Abstracts, 2022
▶ Demo 🗎 Paper

NeuroCartography: Scalable Automatic Visual Summarization of Concepts in Deep Neural Networks
Haekyu Park, Nilaksh Das, Rahul Duggal, Austin P. Wright, Omar Shaikh, Fred Hohman, Duen Horng Chau
IEEE Visualization Conference (VIS), Virtual, 2021
▶ Demo 🗎 Paper

RECAST: Enabling User Recourse and Interpretability of Toxicity Detection Models with Interactive Visualization
Austin P. Wright, Omar Shaikh, Haekyu Park, Will Epperson, Muhammed Ahmed, Stephane Pinel, Duen Horng Chau, Diyi Yang
24th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW), 2021.
🗎 Paper

SkeletonVis: Interactive Visualization for Understanding Adversarial Attacks on Human Action Recognition Models
Haekyu Park, Zijie J. Wang, Nilaksh Das, Anindya S. Paul, Pruthvi Perumalla, Zhiyan Zhou, Duen Horng Chau
AAAI, Demo, Virtual, 2021.
▶ Demo 🗎 Paper

Bluff: Interactively Deciphering Adversarial Attacks on Deep Neural Networks
Nilaksh Das*, Haekyu Park*, Zijie J. Wang, Fred Hohman, Robert Firstman, Emily Rogers, Duen Horng Chau
IEEE Visualization Conference, (VIS), Salt Lake City, UT, USA, 2020.
** Authors contributed equally.*
▶ Demo 🗎 Paper

CNN Explainer: Learning Convolutional Neural Networks with Interactive Visualization
Zijie J. Wang, Robert Turko, Omar Shaikh, Haekyu Park, Nilaksh Das, Fred Hohman, Minsuk Kahng, Duen Horng Chau
IEEE Conference on Visual Analytics Science and Technology, (VAST), Salt Lake City, UT, USA, 2020.
▶ Demo 🗎 Paper

A Comparative Analysis of Industry Human-AI Interaction Guidelines
Austin P. Wright, Zijie J. Wang, Haekyu Park, Grace Guo, Fabian Sperrle, Mennatallah El-Assady, Alex Endert, Daniel Keim, Duen Horng Chau
IEEE Visualization Conference, Workshop on Trust and Expertise in Visual Analytics (TREX), Salt Lake City, UT, USA, 2020.
🗎 Paper

Argo Lite: Open-Source Interactive Graph Exploration and Visualization in Browsers
Siwei Li, Zhiyan Zhou, Anish Upadhayay, Omar Shaikh, Scott Freitas, Haekyu Park, Zijie J. Wang, Susanta Routray, Matthew Hull, Duen Horng Chau
ACM International Conference on Information and Knowledge Management, (CIKM), Resource Track, Online, 2020.
▶ Demo 🗎 Paper

Massif: Interactive Interpretation of Adversarial Attacks on Deep Learning
Nilaksh Das*, Haekyu Park*, Zijie J. Wang, Fred Hohman, Robert Firstman, Emily Rogers, Duen Horng Chau
ACM CHI Conference on Human Factors in Computing Systems (CHI), Late-Breaking Works, Honolulu, Hawaii, USA, 2020.
** Authors contributed equally.*
🗎 Paper

CNN 101: Interactive Visual Learning for Convolutional Neural Networks
Zijie J. Wang, Robert Turko, Omar Shaikh, Haekyu Park, Nilaksh Das, Fred Hohman, Minsuk Kahng, Duen Horng Chau
ACM CHI Conference on Human Factors in Computing Systems (CHI), Late-Breaking Works, Honolulu, Hawaii, USA, 2020.
🗎 Paper

Summit: Scaling Deep Learning Interpretability by Visualizing Activation and Attribution Summarizations
Fred Hohman, Haekyu Park, Caleb Robinson, Duen Horng Chau
IEEE Transactions on Visualization and Computer Graphics (TVCG), Vancouver, BC, Canada, 2020.
▶ Demo 🗎 Paper

Visual Analytics for Interpretability on Deep Neural Networks
Haekyu Park, Fred Hohman, Nilaksh Das, Caleb Robinson, Duen Horng Chau
Women in Machine Learning Workshop (WiML), co-located with NeurIPS 2019, Vancouver, BC, Canada, 2019.

MLsploit: A Framework for Interactive Experimentation with Adversarial Machine Learning Research
Nilaksh Das, Siwei Li, Chanil Jeon, Jinho Jung, Shang-Tse Chen, Carter Yagemann, Evan Downing, Haekyu Park, Evan Yang, Li Chen, Michael Kounavis, Ravi Sahita, David Durham, Scott Buck, Duen Horng Chau, Taesoo Kim, Wenke Lee
ACM SIGKDD Conference on Konwledge Discovery and Data Mining (KDD), KDD Project, Anchorage, Alaska, USA, 2019.
▶ Demo 🗎 Paper

NeuralDivergence: Exploring and Understanding Neural Networks by Comparing Activation Distributions
Haekyu Park, Fred Hohman, Duen Horng Chau
IEEE Pacific Visualization Symposium (PacificVis), Bangkok, Thailand, 2019.
▶ Demo 🗎 Paper

SIDE: Representation Learning in Signed Directed Networks
Junghwan Kim, Haekyu Park, Ji-Eun Lee, U Kang
The Web Conference (Previously known as WWW, World Wide Web Conference), Lyon, France, 2018.
🌐 Webpage 🗎 Paper

A Comparative Study of Matrix Factorization and Random Walk with Restart in Recommender Systems
Haekyu Park, Jinhong Jung, U Kang
IEEE International Conference on Big Data (BigData), Boston, MA, USA, 2017.
🌐 Webpage 🗎 Paper

Open-Source Research Projects

NeuroCartography: Scalable Automatic Visual Summarization of Concepts in Deep Neural Networks 2021
Keywords: Deep Learning Interpretability, Visualization, Human Interpretable Concepts Learned by a Model, Concept Cascade
Interactive visual system that scalably summarizes and visualizes concepts learned by neural networks.
It was published at IEEE Visualization Conference (VIS), 2021.
Haekyu Park, Nilaksh Das, Rahul Duggal, Austin P. Wright, Omar Shaikh, Fred Hohman, Duen Horng Chau
▶ Demo

SkeletonVis: Interactive Visualization for Understanding Adversarial Attacks on Human Action Recognition Models 2021
Keywords: Adversarial Attacks, Human Action Recognition
Interactive visual system for understanding vulnerability of human action recognition model.
It was published at AAAI Demo, 2021.
Haekyu Park, Zijie J. Wang, Nilaksh Das, Anindya S. Paul, Pruthvi Perumalla, Zhiyan Zhou, Duen Horng Chau
▶ Demo

CNN Explainer: Learning Convolutional Neural Networks with Interactive Visualization 2020
Keywords: Deep Learning Education, Interactive Visualization, Interactive Animation
Interactive visual system for learning Convolutional Neural Networks.
It was published at IEEE VIS (VAST, TVCG), 2020.
Zijie Jay Wang, Robert Turko, Omar Shaikh, Haekyu Park, Nilaksh Das, Fred Hohman, Minsuk Kahng, Duen Horng (Polo) Chau
▶ Demo 🏅 Top of Github Trending ★ 4,905 Github stars (as of Oct 2020)

Bluff: Interactively Deciphering Adversarial Attacks on Deep Neural Networks 2020
Keywords: Adversarial Attacks, Neural Network Interpretability, Activation Pathways, Interactive Visual Analytics
Interactive system for visualizing, characterizing, and deciphering adversarial attacks on vision-based neural networks.
It was published at IEEE VIS, 2020.
Nilaksh Das*, Haekyu Park*, Zijie Jay Wang, Fred Hohman, Robert Firstman, Emily Rogers, Duen Horng Chau
(* Equal Contribution)
▶ Demo

Summit: Scaling Deep Learning Interpretability by Visualizing Activation and Attribution Summarizations 2019

Keywords: Neural Network Interpretability, Attribution Graph, Interactive Visual Analytics
Interactive visualization that scalably summarizes what features a deep learning model has learned and how those features interact to make predictions.
It was published at IEEE VIS (VAST, TVCG), 2019.
Fred Hohman, Haekyu Park, Caleb Robinson, Duen Horng Chau
▶ Demo

MLsploit: A Framework for Interactive Experimentation with Adversarial Machine Learning Research 2019

Keywords: Adversarial Attacks and Defenses for Machine Learning Models, Interactive Experimentation
User-friendly, cloud-based system that enables researchers and practitioners to rapidly evaluate and compare state-of-the-art adversarial attacks and defenses for machine learning (ML) models.
It was published at a KDD 2019 Project Showcase.
▶ Demo

SIDE: Representation Learning in Signed Directed Networks 2018

Keywords: Network Embedding, Signed Weighted Directed Graph
General network embedding method that represents both sign and direction of edges in the embedding space.
It was published at the Web Conference (WWW), 2018.
🌐 Webpage

A Comparative Study of Matrix Factorization and Random Walk with Restart in Recommender Systems 2017

Keywords: Recommender System, Matrix Factorization (MF), Random Walk with Restart (RWR)
We provide a comparative study of MF and RWR, which are the most representative methods for recommender systems.
It was published at IEEE Big Data, 2017.
🌐 Webpage

Other Projects

Accelerated Data Science Teaching Kit for Educators 2021

Keywords: GPU-accelerated Data Science, RAPIDS, NVIDIA Teaching Kits
The first version of its GPU Accelerated Data Science Teaching Kit for educators.
Presented at NVIDIA's Graphics Technology Conference (GTC) 2021: Bridging Data Analytics and Machine Learning Skill Gaps with RAPIDS and the New Accelerated Data Science Teaching Kit for University Educators [S31763]
🌐 Data Science Teaching Kit 🌐 Blog

DARPA Guaranteeing AI Robustness against Deception (GARD) 2020-2021

Keywords: Defenses for Adversarial Examples, Robustness, Defense using Semantic Coherence
We develop defenses for adversarial attacks on object detector for both RGB images and single-camera video. We augment this object detector to support spatial, temporal, semantic coherence in videos.

RAPIDS and Cybersecurity: A Network Use Case 2019

Keywords: RAPIDS, NVIDIA, GPU-acceleration, Graph, Personalized Page Rank
We showcase an approach to flagging anomalous network communications in a large graph using a combination of structural graph features and graph analytics, running end-to-end in RAPIDS.
Presented at cybersecurity use case notebook.

Recommender System for Videos on Oksusu Application 2017

Keywords: Deep Learning, Sequence/Word Embedding, Approx. k-NN, Heterogeneous Features
Our system recommends videos to users of Oksusu application, handling massive data on users' behaviors and heterogeneous information of videos.
SK Telecom, Seoul, Republic of Korea

A Fast Data Compression with Shared Virtual Memory in Heterogeneous System Architecture 2017
Keywords: OpenCL, GPGPU, SVM, HSA
I used general purpose computing on graphics processing units (GPGPU) and Shared Virtual Memory (SVM) in Heterogeneous System Architecture (HSA) for fast data deduplication methods. GPGPU and HSA provide a powerful basis for parallel computing in an easy programmable and efficient way.
Undergraduate thesis

Personalized Recommendation for Credit Card Rewards 2016
Keywords: Coupled Matrix Factorization, Time Series Data
We provide personalized recommendations for credit card rewards to customers using various side information of users and items. The main algorithm is TCMF (Time Coupled Matrix Factorization).
Hyundai Card, Seoul, Republic of Korea
📰 News article (in Korean)

Talks

NeuroCartography: Scalable Automatic Visual Summarization of Concepts in Deep Neural Networks
Haekyu Park, Nilaksh Das, Rahul Duggal, Austin P. Wright, Omar Shaikh, Fred Hohman, Duen Horng Chau
Oct 2021, IEEE Visualization Conference (VIS)

SkeletonVis: Interactive Visualization for Understanding Adversarial Attacks on Human Action Recognition Models
Haekyu Park, Zijie Jay Wang, Nilaksh Das, Anindya S. Paul, Pruthvi Perumalla, Zhiyan Zhou, Duen Horng Chau
Feb 2021, Poster Presentation, AAAI

Bluff: Interactively Deciphering Adversarial Attacks on Deep Neural Networks
Nilaksh Das*, Haekyu Park*, Zijie Jay Wang, Fred Hohman, Robert Firstman, Emily Rogers, Duen Horng Chau
(* Equal Contribution)
Oct 2020, Oral Presentation, IEEE VIS
Oct 2020, Presentation, Michigan Institute for Data Science (MIDAS) Consortium for researchers in Training

Accelerated Data Science in the Classroom: Teaching Analytics and Machine Learning with RAPIDS
Polo Chau and Haekyu Park
Mar 2020, Talk, NVIDIA's GPU Technology Conference (GTC)

NeuralDivergence: Exploring and Understanding Neural Networks by Comparing Activation Distributions
Apr 2019, Poster Presentation, PacificVis

A Comparative Study of Matrix Factorization and Random Walk with Restart in Recommender Systems
Dec 2017, Oral Presentation, IEEE Big Data

Tutorial

RAPIDS and Cybersecurity: A Network Use Case
Keywords: RAPIDS, NVIDIA, GPU-acceleration, Graph, Personalized Page Rank
Presented at KDD 2019 NVIDIA RAPIDS tutorial with the cybersecurity use case notebook

Teaching

Graduate Teaching Assistant
Georgia Institute of Technology, Atlanta, GA

Data and Visual Analytics (CSE 6242)
Fall 2019, Fall 2021
Instructor: Polo Chau

Mentoring

Aiswaya Bhagavatula 2021
M.S. in Computational Science and Engineering, Georgia Institute of Technology
GPU accelerated data science teaching kit
AI Robustness against Adversarial Attacks

Sushanto Praharaj 2021
M.S. in Computational Science and Engineering, Georgia Institute of Technology
AI Robustness against Adversarial Attacks
Received Marshall D. Williamson Fellowship award

Jon Saad-Falcon 2021
B.S./M.S. in Computer Science, Georgia Institute of Technology
GPU accelerated data science teaching kit
Received Donald V. Jackson Fellowship award

Kevin Li 2021
B.S. in Computer Science, Georgia Institute of Technology
GPU accelerated data science teaching kit

Zhiyan Zhou 2021
B.S. in Computer Science, Georgia Institute of Technology
AI Robustness against Adversarial Attacks

Megan Dass 2021
B.S. in Computer Science, Georgia Institute of Technology
AI Robustness against Adversarial Attacks
Received Outstanding Freshman Award

Omar Shaikh 2019-2020
B.S. in Computer Science, Georgia Institute of Technology
Visualization for natural language processing
Received Outstanding Freshman Award
Received Sigma Xi Best Undergraduate Research Award

Rob Firstman 2019-2020
B.S. in Computer Science, Georgia Institute of Technology
Visualization for deep learning interpretability

Robert Turko 2019-2020
B.S. in Computer Science, Georgia Institute of Technology
Visualization for machine learning education
Received Outstanding Senior Award

Licenses and Certifications

Licenses and Certifications
NVIDIA DLI Certificate – DLI Platform Course for Instructors, NVIDIA Deep Learning Institute
NVIDIA DLI Certificate – Fundamentals of Deep Learning for Computer Vision, NVIDIA Deep Learning Institute

Technical Skills

Programming Languages
Python, JavaScript, TypeScript, HTML, R, Matlab, Java, C, C++, Ocaml, Scheme

Machine Learning / Deep Learning / Data Science
TensorFlow, PyTorch, Keras, scikit-learn, OpenCV, Numpy, Pandas, SciPy, NetworkX

GPU-accelerated Data Science
cuGraph, cuDF, cuML, BlazingSQL, OpenCL

Interface / Data Visualization
React, D3.js, Three.js, WebGL, HoloViews, Matplotlib, WebGL, ggplot