Hamza Tahboub

linkedin.com/in/hamzatahboub
Availability: Summer, 2025

tahboub.h@northeastern.edu
+1 (858) 371-8866
Boston, MA

Education

- **Northeastern University, Khoury College of Computer Sciences** — Boston, United States
  Bachelor of Science — Computer Science and Mathematics; GPA: 3.95, Dean's List — *2021 – 2025*

  Relevant Courses: Practical Neural Networks, Adv. Linear Algebra, Statistics and Stochastic Processes, Adv. Programming with Data, Data Management and Processing (graduate), Number Theory, Programming in C++, Matrix Methods in ML, Technical Writing

Skills

- **Languages**: Python, Java, C++, SQL, LaTeX, Lisp, MATLAB, C#
- **Tools**: PyTorch, Git, Apache Spark, Hadoop, NumPy, Pandas, Sklearn, Matplotlib, MongoDB, Azure, GCP, AWS

Experience

- **Professor Huaizu Jiang's Visual Intelligence Lab — Northeastern University** — Boston, MA
  Research Assistant — *Aug 2022 – Present*
  - Leading a computer vision project on social interaction modeling from an egocentric/first-person perspective. Aiming to submit a paper on our competitive agent in social interaction games to ICCV 2025.
  - Designed and evaluated multimodal transformer-based architectures to leverage the unique advantages of our egocentric data (additional information such as what/who each agent is looking at) while mitigating disadvantages (such as the shakiness of the viewpoint).
  - Devised and formulated new loss functions to help the agent base its responses on strategic thinking, which work by pulling together the embeddings of utterances that should have the same strategies.
  - Received the Summit undergraduate research award in support of my work, including funding.
  - Assisted on other projects focused on visual commonsense reasoning abilities.

- **Genentech, Subsidiary of Roche** — San Francisco, CA
  Natural Language Processing Co-op — *July 2023 – Dec 2023*
  - Contributed to experimental medical NLP research, designing experiments and reviewing new methods.
  - Distilled capabilities from attention-based language models to smaller ones while maintaining accuracy, reducing long-term computation costs by over 95% and reducing reliance on closed-source models.
  - Curated a synthetic dataset of over 100k samples for training transformers and other deep learning models for specialized medical QA.
  - Implemented models from papers for the computer vision team for cellular semantic segmentation.
  - Developed an embedding-based semantic search engine to retrieve from medical corpora of 150k+ documents.

- **Khoury College of Computer Sciences — Northeastern University** — Boston, MA
  Teaching Assistant — *Jan 2022 – July 2022*
  - Mentored over 80 students for two semesters as a TA for the Fundamentals of Computer Science course.
  - Provided guidance in office hours, graded assignments, and led labs in which students practiced new material.

Projects

- **MarkovPatch: Random Image Masks for Interpretable AI** — *Fall 2022*
  - Applied image masks to a pre-trained classification deep neural network during inference to identify spatial features of significance.
  - Developed a random mask generator by sampling a second-order Markov chain. The distribution parameters were adjusted to alter the size and spatial correlation of the masks' patches.
  - Applying this method to a cat classifier, demonstrated that the model is paying more attention to contour features and specific feline attributes like ears and whiskers.

- **Assigning TAs to Labs Using Evolutionary Computing** — *Fall 2022*
  - Formulated matching teaching assistants to lab sessions with constraints as a cost optimization problem.
  - Developed a program that applies evolutionary computing principles to search for the minimum-cost solution.
  - Wrote scoring functions to quantify progress and agents that "mutated" solutions to search for better ones.
  - Was able to quickly find optimal solutions 100% of the time for realistic parameters using my method.

Certifications

- **Deep Learning Specialization**: Neural Networks and Deep Learning Fundamentals, Hyperparameter Tuning, Regularization and Optimization, Structuring Machine Learning Projects, Convolutional Neural Networks, Sequence Models. (online course taught by Professor Andrew Ng)

Interests

- Running — Road cycling — Math — Computer graphics — Linguistics and language learning — Skiing