

Amin Karimi Monsefi

PH.D. STUDENT OF COMPUTER SCIENCE.

Columbus, Ohio, USA

☎ (+1) 740-803-4733 | ✉ amin.karimi1992@gmail.com | 🏠 7amin.github.io | 🌐 amin-karimi-monsefi-5672b4153

“Be the change that you want to see in the world.”

Summary

Amin Karimi Monsefi is a dedicated Ph.D. student in Computer Science at The Ohio State University, specializing in **Computer Vision, Generative Models, Self-Supervised Learning, and Diffusion Models** under the guidance of Professor Rajiv Ramnath. His research advances the field through innovative **image and video generation** methods, including projects like Multi-Guided Image Inpainting and Multi-Modal Conditional Video Generation. Amin explores how the creativity of Large Language Models (LLMs) can be leveraged in video generation using diffusion models. He also designs self-supervised learning approaches to extract meaningful representations from unlabeled data, applying these techniques to both general vision tasks and medical image analysis—where he developed Masked LoGoNet for efficient 3D image segmentation. With **18 papers accepted or under review** at prestigious **conferences such as CVPR, ICCV, ICLR, and SIGKDD**, he is also a **reviewer for conferences like SIGKDD (2024–2025), ICLR 2025, CVPR 2025, and WACV 2025**, reflecting his commitment to advancing computer vision and contributing significantly to image and video generation technologies.

Education

PhD. in Computer Science

Columbus, Ohio, USA

THE OHIO STATE UNIVERSITY

Jan. 2022 - Exp. Present

- **Supervisor:** Prof. Rajiv Ramnath
- **Image and Video Generation:** Developing innovative methods for generating high-quality images and videos. Projects include Multi-Guided Image Inpainting and Multi-Modal Conditional Video Generation. Exploring how the creativity of Large Language Models (LLMs) can be utilized in video generation with diffusion models.
- **Multi-Modal Conditional Video Generation:** Focused on harnessing diverse conditions to enhance visual synthesis, this project aims to generate videos that are not only visually appealing but also contextually rich and dynamic. By incorporating various modalities and conditions, the project seeks to create more immersive and informative video content.
- **Self-Supervised Approach for General Images:** This project involves the exploration of a self-supervised approach tailored for general images. It leverages multimodal architectures such as CLIP to enhance the image encoder and employs Data-centric AI approaches to build a better encoder. The aim is to establish robust foundation models for various vision-related tasks, including image generation and classification.
- **Medical Image Analysis:** Focused on addressing the challenge of limited labeled data, this project adopts a self-supervised learning approach. By leveraging this technique, it trains models using unlabeled medical images, allowing them to learn meaningful representations and extract valuable features from the data, thus enhancing the analysis and interpretation of medical images.
- **Smart Mobility Research:** This project centers on employing deep learning methods for time series analysis within the field of smart mobility. The focus is crafting models to forecast and analyze the likelihood of accidents and road construction in designated zones. The project introduces a multi-source dataset and a multimodal architecture, harnessing diverse data streams such as historical records, weather data, satellite and map imagery, and demographic details to advance the field.

M.Sc. in Computer Engineering with focus on Software

Tehran, Iran

SHAHID BEHESHTI UNIVERSITY

Sep. 2015 - Exp. Sep. 2017

- **Thesis:** An Approach for Automatic Software Test Data Generation Using Machine Learning and Program Static Structure
- **Supervisor:** Prof. H.Haghighi
- I achieved **Ranked 1st**, National M.Sc Entrance Exam of Computer Engineering, Among More Than 20,000 Students

B.Sc. in Computer Engineering with focus on Hardware

Tehran, Iran

SHAHID BEHESHTI UNIVERSITY

Sep. 2011 - Exp. Sep. 2015

- **Thesis:** Applying Reinforcement Learning on Multi-Agent Environments
- **Supervisor:** Prof. M. Abdoos

Papers

Enhancing Open Vocabulary Segmentation through Semantic Attention and Path-Aligned Contrastive Learning

Under Submission

AMIN KARIMI MONSEFI, KISHORE PRAKASH SAILAJA, WEI-LUN (HARRY) CHAO, RAJIV RAMNATH

2025

Contrastive Generation: A Cutting-Edge Approach to Image Generation

Under Submission

AMIN KARIMI MONSEFI, KISHORE PRAKASH SAILAJA, RAJIV RAMNATH, SER-NAM LIM

2025

TaxaDiffusion: Progressively Trained Diffusion Model for Fine-Grained Species Generation

ICCV

AMIN KARIMI MONSEFI, MRIDUL KHURANA, RAJIV RAMNATH, ANUJ KARPATNE, WEI-LUN (HARRY) CHAO, CHENG ZHANG

2025

- International Conference on Computer Vision
- Honolulu, Hawai'i, USA

KnobGen: Controlling the Sophistication of Artwork in Sketch-Based Diffusion Models

CVEU Workshop - CVPR

AMIN KARIMI MONSEFI, POUYAN BORESHNAVAR, MENGXI ZHOU, WEI-LUN (HARRY) CHAO, ALPER YILMAZ, RAJIV RAMNATH

2025

- CVPR - AI for Creative Visual Content Generation Editing and Understanding (CVEU) Workshop
- Nashville, Tennessee, USA

Frequency-Guided Masking for Enhanced Vision Self-Supervised Learning

ICLR

AMIN KARIMI MONSEFI, MENGXI ZHOU, NASTARAN KARIMI MONSEFI, SERNAM LIM, WEI-LUN (HARRY) CHAO, RAJIV RAMNATH

2025

- The Thirteenth International Conference on Learning Representations
- Singapore

DetailCLIP: Detail-Oriented CLIP for Fine-Grained Tasks

SSI-FM Workshop - ICLR

AMIN KARIMI MONSEFI, KISHORE PRAKASH SAILAJA, ALI ALILOOEE, SER-NAM LIM, RAJIV RAMNATH

2025

- ICLR - Scaling Self-Improving Foundation Models (SSI - FM) Workshop
- Singapore

Masked LoGoNet: Fast and Accurate 3D Image Analysis for Medical Domain

30TH ACM SIGKDD - KDD

AMIN KARIMI MONSEFI, PAYAM KARISANI, MENGXI ZHOU, STACEY CHOI, NATHAN DOBLE, HENG JI, SRINIVASAN

2024

PARTHASARATHY, RAJIV RAMNATH

- Conference On Knowledge Discovery And Data Mining
- Barcelona, Spain

CrashFormer: A Multimodal Architecture to Predict the Risk of Crash

31st ACM SIGSPATIAL

AMIN KARIMI MONSEFI, POUYA SHIRI, AHMAD MOHAMMADSHIRAZI, NASTARAN KARIMI MONSEFI, RON DAVIES, SOBHAN

2023

MOOSAVI, RAJIV RAMNATH

- UrbanAI '23: Proceedings of the 1st ACM SIGSPATIAL
- Hamburg, Germany

Novel Physics-Based Machine-Learning Models for Indoor Air Quality Approximations

29TH ACM SIGKDD - KDD

AHMAD MOHAMMADSHIRAZI, AIDA NADAFIAN, AMIN KARIMI MONSEFI, MOHAMMAD H RAFIEI, RAJIV RAMNATH

2023

- Conference On Knowledge Discovery And Data Mining
- Long Beach, California, USA

Smart and collaborative industrial IoT: A federated learning and data space approach

Elsevier - Journal - Impact Factor 7.5

BAHAR FARAHANI AND AMIN KARIMI MONSEFI

2023

- Journal: Digital Communications and Networks

Will there be a construction? Predicting road constructions based on heterogeneous spatiotemporal data

30th ACM SIGSPATIAL

AMIN KARIMI MONSEFI, SOBHAN MOOSAVI AND RAJIV RAMNATH

2022

- Conference: International Conference on Advances in Geographic Information Systems
- Seattle, Washington, USA

Professional Experiences

Apple

Seattle, WA, USA

ML RESEARCH INTERN

May 2025 - Present

- Apple MIND Team
- Working on multimodal vision-language models and efficient training for generative applications.

Higharc

North Carolina, USA

MACHINE LEARNING INTERN

May 2024 - Aug 2024

- Conducting research on semantic and panoptic segmentation tasks.
- Utilizing unlabeled data to pre-train a DETR-based model and addressing challenges of limited labeled data with Self-Supervised learning.
- Implementing domain adaptation approaches to generalize models across multiple datasets with distinct distributions. Also, developing strategies to transfer a trained model from one domain to another effectively.

JIBB

San Francisco, USA

SENIOR MACHINE LEARNING ENGINEER

Dec 2020 - Dec 2021

- Detected and removed unwanted objects from images and videos using advanced object detection techniques.
- Implemented methods to detect colors, remove shadows, and eliminate reflections from images and videos.
- Developed and optimized various U-Net-based architectures to enhance image and video processing tasks.
- Utilized machine learning algorithms and computer vision techniques for real-time image and video enhancement.

BlueBitSoft

Tehran, Iran

CTO

Dec 2018 - Dec 2021

- Served as the technical consultant, providing strategic guidance and oversight for technical projects.
- Designed high-level architecture for pharmacy software solutions, ensuring scalability, reliability, and efficiency.
- Collaborated with cross-functional teams, including software developers and business experts, to align technical strategies with business goals.
- Implemented best practices in software engineering, including agile methodologies and continuous integration/continuous deployment (CI/CD) pipelines.
- Led efforts to enhance software performance and security, ensuring compliance with industry standards and regulations.

TAPSI (Online Taxi)

Tehran, Iran

SENIOR DATA SCIENTIST & BACK-END DEVELOPER

Mar. 2018 - Dec 2020

- Designed and implemented pricing micro-services using Python, communicating through event messaging with RabbitMQ.
- Developed a fake GPS detection algorithm to prevent fraud and have a safe and secure system.
- Created intelligent backend features based on passengers' data, including origin suggestion, destination suggestion, favorite location suggestion, and ride suggestion using unsupervised.
- Implemented a micro-service for estimated arrival time (ETA) using drivers' GPS information. Introduced a new algorithm for ETA based on GPS data and published a paper on this work. I also published a paper in this manner.
- Developed a smart map for predicting locations with higher ride demand rates in cities.
- Utilized Python and Node.js for back-end development, Spark, Kafka, and R for analytical purposes, and PostgreSQL, MongoDB, and Redis.
- Python, NodeJS, Spark, Kafka, RabbitMQ, PostgreSQL and MongoDB

Honors & Awards

REVIEWER SERVICE RECOGNITION

KDD 2025 **Outstanding Reviewer**, Second round - 10% of top reviewers received.

Toronto, Canada

KDD 2025 **Excellent Reviewer**, First round - 20% of top reviewers received.

Toronto, Canada

ACM-ICPC COACH AND TEAM LEAD

Apr 2018 **Coach**, 42nd Annual World Finals

Beijing, China

Dec 2017 **Coach**, 7th and 17th teams, Regional Contest

Tehran, Iran

Nov 2017 **Coach**, 10th and 16th teams, Internet Programming Contest

Tehran, Iran

Oct 2017 **Coach**, The 1st team, Regional Contest

Kabul, Afghanistan

Oct 2017 **Coach**, The 1st team, Online Programming Contest

Kabul, Afghanistan

Oct 2015 **Coach**, The 11th team, Iran-Internet-PC-IU Online Programming Contest

Tehran, Iran

Oct 2015 **Coach**, The 23rd team, Iran-Internet-PC-SBU Online Programming Contest

Tehran, Iran

Nov 2014 **Coach**, The 12th, 53rd, and 131st teams, Internet Programming Contest

Tehran, Iran

ACM-ICPC CONTESTANT

Dec 2014 **10th Place**, Asia Regional Contest

Tehran, Iran

Oct 2014 **2nd Place**, Qualification Contest, Shahid Beheshti University

Tehran, Iran

Dec 2013 **3rd Place**, Asia Regional Contest

Tehran, Iran

Oct 2013 **1st Place**, Qualification Contest, Shahid Beheshti University

Tehran, Iran

May 2013 **5th Place**, The 5th ACM National Collegiate Programming Contest, University of Kashan

Kashan, Iran

Apr 2013 **8th Place**, The 1st ACM National Collegiate Programming Contest, University of Shahrood

Shahrood, Iran

Dec 2012 **25th Place**, Asia Regional Contest

Tehran, Iran

Nov 2012 **16th Place**, The 3rd ACM National Collegiate Programming Contest, Amir Kabir University of Technology

Tehran, Iran

Oct 2012 **2nd Place**, Qualification Contest, Shahid Beheshti University

Tehran, Iran

OTHER

June 2015 **Ranked 1st**, National M.Sc Entrance Exam of Computer Engineering, Among More Than 10.000 Students

Iran

June 2011 **Ranked Top 1%**, National B.Sc Entrance Exam of Mathematics, Among More Than 256,000 Students

Iran

Mar 2009 **Bronze Medal**, Mathematics Olympiad of Waterloo University

Tehran - Waterloo

Teaching Experiences

TEACHING ASSISTANT

Spr 2025	Computer Vision for Human-Computer Interaction , Teacher: Dr. Wei-Lun (Harry) Chao	Ohio State Uni
5 times	Modeling and Problem Solving with Spreadsheets and Databases , Fall 22/23/25 and Spring 23/24	Ohio State Uni
Spr 2012	Introduction to Programming , Teacher: Dr. A. M. Aznaveh	Shahid Beheshti Uni
Fall 2012	Introduction to Programming , Teacher: Dr. A. M. Aznaveh	Shahid Beheshti Uni
Spr 2013	Introduction to Programming , Teacher: Dr. M. H. Moayeri	Shahid Beheshti Uni
Spr 2013	Discrete Mathematics , Teacher: Dr. F. Safaie	Shahid Beheshti Uni
Fall 2013	Discrete Mathematics , Teacher: Dr. H. Haghighi	Shahid Beheshti Uni
Spr 2014	Discrete Mathematics , Teacher: Dr. F. Safaie	Shahid Beheshti Uni
Spr 2015	Introduction to Algorithm , Teacher: Dr. E. Bavafaye	Shahid Beheshti Uni
Fall 2015	Data Structure , Teacher: Dr. M. E. Moghadam	Shahid Beheshti Uni
Fall 2015	Artificial Intelligence , Teacher: Dr. M. Shamsfard	Shahid Beheshti Uni
Fall 2015	Introduction to Algorithm , Teacher: Dr. R. Ghavamizade	Shahid Beheshti Uni
Fall 2016	Data Structure , Teacher: Dr. M. E. Moghadam	Shahid Beheshti Uni
Fall 2016	Artificial Intelligence , Teacher: Dr. M. Shamsfard	Shahid Beheshti Uni
Spr 2017	Discrete Mathematics , Teacher: Dr. H. Haghighi	Shahid Beheshti Uni
Spr 2018	Discrete Mathematics , Teacher: Dr. H. Haghighi	Shahid Beheshti Uni