Last Updated on Oct 5th, 2024

Sara Ghazanfari

github.com/SaraGhazanfari saraghazanfari.github.io linkedin.com/in/sara-ghazanfari sg7457@nyu.edu

Education

Ph.D. Candidate Jan. 2023 - Jan. 2027
Electrical and Computer Engineering Department, New York University (NYU) GPA: 3.95/4.0
Advisor: Siddharth Garg, Farshad Khorrami
Research Area: Deep Learning, Adversarial Machine Learning, Computer Vision
Courses: Probability & Stochastic (A), Deep Learning (A), Algorithmic Machine Learning & Data Science (A^-), Linear Systems (A), Machine Learning for Cybersecurity (A).

Master of Science Sep. 2018 - July 2021
Computer Engineering Department, Sharif University of Technology (SUT) GPA: 4.0/4.0

Bachelor of Science Sep. 2013 - Feb. 2018
Computer Engineering Department, Sharif University of Technology (SUT) GPA: 3.6/4.0

Publications

EMMA: Efficient Visual Alignment in Multi-Modal LLMs Submitted to ICLR 2025
LipSim: A Provably Robust Perceptual Similarity Metric ICLR 2024
R-LPIPS: An Adversarially Robust Perceptual Similarity Metric ICML Workshop 2023

Skills

Programming Languages: Python, C, Java, R.
Machine Learning Tools: Scikit-Learn, Spacy, Pytorch.
Database: PostgreSQL, Redis, Elasticsearch.
Software and Tools: Microsoft Office, MATLAB, Jupyter Notebook, PyCharm, Git/GitHub, Unix Shell, Kafka.
Web frameworks: Django, Flask.

Teaching Experiences

Head TA, Machine Learning Feb. 2021 - Aug. 2021
Instructor: Prof. Soleymani, Prof. Sharifi-Zarchi at SUT.

- Coordinating all course plans with the teaching group.
- Devising assignments in collaboration with TAs.

TA, Modern Information Retrieval Feb. 2020 - Aug. 2020
Instructor: Prof. Soleymani at SUT.
Description: Designing the course project.

TA, Machine Learning Feb. 2019 - Aug. 2019
Instructor: Prof. Rohban, Prof. Sharifi-Zarchi at SUT.
Description: Designing and checking one assignment for the course.

Research Interests

Large Multi-modal Models, Computer Vision, Trustworthy Machine Learning.

Projects

EMMA: Efficient Visual Alignment in Multi-Modal LLMs

In this work, we propose EMMA (Efficient Multi-Modal Adaptation), a lightweight cross-modality module designed to efficiently fuse visual and textual encodings, generating instruction-aware visual representations for the language model. Our key contributions include: (1) an efficient early fusion mechanism that integrates vision and language representations with minimal added parameters (less than 0.2% increase in model size), (2) an in-depth interpretability analysis that sheds light on the internal mechanisms of the proposed method; (3) comprehensive experiments that demonstrate notable improvements on both specialized and general benchmarks for MLLMs. Empirical results show that EMMA boosts performance across multiple tasks by up to 9.3% while significantly improving robustness against hallucinations.

LipSim: A Provably Robust Perceptual Similarity Metric

In this work, we demonstrate the vulnerability of the SOTA perceptual similarity metric based on an ensemble of ViT-based feature extractors to adversarial attacks. We then propose a framework to train a robust perceptual similarity metric called LipSim (Lipschitz Similarity Metric) with provable guarantees by leveraging 1-Lipschitz neural networks as backbone and knowledge distillation approach to distill the knowledge of the SOTA models. Finally, a comprehensive set of experiments shows the performance of LipSim in terms of natural and certified scores and on the image retrieval application.

R-LPIPS: An Adversarially Robust Perceptual Similarity Metric

In this work, we show that the LPIPS metric is sensitive to adversarial perturbation and propose the use of Adversarial Training to build a new Robust Learned Perceptual Image Patch Similarity (R-LPIPS) that leverages adversarially trained deep features. Based on an adversarial evaluation, we demonstrate the robustness of R-LPIPS to adversarial examples compared to the LPIPS metric. Finally, we showed that the perceptual defense achieved over LPIPS metrics could easily be broken by stronger attacks developed based on R-LPIPS.

FaRS: Fast Randomized Smoothing

In this project, we leverage the Lipschitz bound of the 1-Lipschitz networks as the backbone of our model and add the linear layer for the classification, and therefore we show that the Monte Carlo sampling is only required for the Linear layer and not for whole model including the backbone.

RoDINO: Boosting Empirical Robustness of Representations by Leveraging Modern Attacks

In this project, we propose RoDINO (Robust DINO) which is a method to boost the empirical robustness of downstream tasks by leveraging PGD attack to generate adversary images and adversarially train DINO which is a self-supervised representation learning model with Vision Transformers backbone.

Adversarial Attacks against the FixCaps Model for Skin Cancer Detection

In this project, we reproduce the accuracy of the FixCaps model on the HAM10000 dataset and explore the robustness of FixCaps to three extensively used attacks, FGSM, PGD, and UAP.

Deep Learning Course Mini-project

In this project a Resnet network with at most 5 million parameters is trained.

Experience

Research Assistant, New York University (NYU), US — Jan. 2023 - Jan. 2027

Electrical and Computer Engineering Department

My research is focused on building robust and scalable multimodal perception systems that can operate in the real world.

- My latest work, **EMMA**, proposes an efficient modality adaptation module to provide **instruction-aware visual representations to MLLM** (Multi-modal Large Language Models).
- My earlier work **LipSim** introduces a **robust perceptual metric** by leveraging the Lipschitz Model as the backbone and distilling from the SOTA perceptual metric.

Technical Team Lead, Narvan Startup Studio, Iran Oct. 2021 - Jan. 2023

The technical team leader of a software development team. The goal was to develop a cryptocurrency exchange website and mobile application. I was in charge of developing the backend of the software with a team of backend developers. Our stack was composed of Python, Django, PostgreSQL, Sentry, and gRPC and we followed the microservices architecture pattern for our architecture.

Data Engineer, Narvan Startup Studio, Iran Aug. 2020 - Oct. 2021

- Designing and implementing data pipelines using Kafka.
- Storing and retrieving data in SQL-like databases such as PostgreSQL and NoSQL databases such as Elasticsearch, InfluxDB, and Redis.
- Designing an NLP pipeline, applied on Persian News.
- Designing and implementing preprocessing components named: Tokenizer, Lemmatize, and Normalizer with Python.
- Implementing and tuning CNN with residual connections for topic modeling (text classification) with Pytorch.

All tasks are done with **Python** programming language.