KIMIA AFSHARI

Computer Vision Engineer

Phone:
Email:
LinkedIn: Linkedin.com/in/kimia-afshari
Location: Santa Barbara, CA
Work Eligibility: U.S. Permanent Resident

PROFESSIONAL SUMMARY

As a Computer Vision engineer, I specialize in designing, developing, and evaluating multi-modal models, and processing large-scale image and video data. My expertise encompasses applying machine learning to computer vision and vision-language systems across various domains. With a robust background in foundational models, deep learning, and techniques for image and video detection, recognition, and tracking, I focus on enhancing autonomous systems. I have hands-on experience with deep learning and distributed data parallel frameworks such as PyTorch, CUDA, and TensorFlow as well as Git, Docker and programming languages C/C++ and Python. Driven by a passion for continuous learning and innovation, I tackle complex real-world challenges with efficient solutions. Beyond my technical skills, I excel in cross-functional teamwork, effective communication, and translating program requirements into reliable solutions.

EDUCATION

Master of Science in Computer Science (GPA: 4.0/4.0)

Sep 2022 – July 2024

University of California, Santa Barbara - Santa Barbara, CA

- *Member of Vision Research Lab (VRL).*
- *Graduate Teaching Assistant in Introduction to Computer Vision (CS181), Problem Solving I - C++ (CS16), and Introduction to Computer Science - Python (CS8) courses.*

Bachelor of Science in Computer Engineering (GPA: 3.64/4.0)

Sep 2016 – Sep 2020

K. N. Toosi University of Technology (KNTU) - Tehran, Iran

- *Member of Machine Vision and Medical Image Processing (MVMIP) and Robotics and Automated Systems (ARAS) Labs.*
- *Teaching assistant in Advanced Programming (Java), Fundamentals of Programming (C++), and Database courses.*

EXPERIENCE

Computer Vision / Software Engineer

July 2024 – Present

Mayachitra Inc. – Santa Barbara, CA

- *Collaborating on a 3D scene reconstruction project to build a 3D model from a limited set of images.*
- *Registering overlapping images taken from different cameras within the same scene.*
- *Employing SuperPoint and RoMa neural models for efficient key points extraction and robust feature matching respectively.*
- *Projecting matching pairs into a global coordinate system and estimating the position of each camera.*
- *Implementing custom techniques to enhance the number of matching pairs and registered images.*
- *Evaluating the pipeline using various parameters and techniques.*
- *Developing and integrating the pipeline within a Docker container to ensure consistency and portability.*

Computer Vision Engineer
Sep 2023 – July 2024

Vision Research Lab (VRL) – Santa Barbara, CA

- *Collaborated on Complex Activity Recognition project, pioneering an advanced surveillance system that incorporates multi-camera person tracking and re-identification to the generated scene graph to understand compositional human actions in multi-view videos.*
- *Researched and evaluated recent single- and multi-camera multi-object trackers and re-identifiers on our challenging and dynamic dataset.*
- *Developed a generalized multi-camera person tracking framework that consistently tracks and re-identifies people across both overlapping and non-overlapping cameras.*
- *Combined appearance- and motion-based features to continuously associate people.*
- *Used hierarchical clustering methods to merge tracklets of the same identity into a unique group.*
- *Addressed challenges related to variance of illumination, frame quality, and occlusion by meticulous feature engineering and tracklet refinement techniques.*
- *Implemented efficient parallelized data loading method that effectively distribute data and optimize model on multiple GPUs, achieving synchronized and balanced workloads that significantly accelerated the training pipeline.*
- *Utilized PyTorch, OpenCV and different deep learning and data manipulation frameworks to implement the pipeline.*

Computer Vision / Software Engineer Intern
July 2023– Sep 2023

SRI International (PARC) – Menlo Park, CA

- *Engineered an Interactive Segmentation Tool specifically for X-ray Tomography data, streamlining the data labeling process.*
- *Innovated a 3D Adaptive Brush to segment regions volumetrically, requiring minimal interactions for comprehensive segmentation.*
- *Employed various machine learning algorithms to introduce versatile brushing techniques tailored for different regions within the volume.*
- *Integrated the Random Forest algorithm to refine segmentation automatically, enhancing the accuracy by filling in the sparse brushings.*
- *Seamlessly incorporated these advanced tools into the Napari viewer platform, significantly improving user experience.*
- *Adopted human-in-the-loop methodologies, increasing the system's robustness and accuracy.*
- *Continuously consulting with clinical experts to ensure that requirements are being met.*

Software Engineer
Sep 2019 – Mar 2022

MVMIP Lab K. N. Toosi University of Technology - Tehran, Iran

- *Developed advanced medical software designed to assist clinics and hospitals in analyzing and processing CT images*
- *Integrated an IPython console, enabling manual function execution and facilitating batch processing capabilities.*
- *Engineered the software with a flexible architecture, allowing for the integration of additional extensions.*
- *Incorporated lung segmentation and classification modules as internal plugins, augmenting its diagnostic capabilities.*
- *Enhanced the software's user interface by utilizing the PyQt framework, offering the end-users an intuitive GUI.*
- *Managed the synchronization and data-sharing processes, ensuring seamless integration and data flow.*
- *Fostered a collaborative work environment by effectively teaming up with cross-functional team members.*

RECENT PROJECTS

Person Tracking and Re-Identification in Distributed Camera Network (M.S. Project)

This project introduces an innovative framework designed to improve surveillance, monitoring and anomaly detection systems through consistent tracking and re-identification of people across both overlapping and non-overlapping cameras. To continuously track and re-identify individuals, a fusion of people's motion and appearance features are leveraged. Addressing the difficulties posed by fluctuating lighting conditions, quality of frames, and occlusions, this approach refines

and connects the tracking results from individual cameras to produce reliable and continuous tracklets. Lastly, hierarchical clustering organizes single-camera tracklets into distinct groups of unique identities.
Deep Learning, Transformer, Multi-camera, Tracking, Clustering, Kalman Filter, Video Recognition, PyTorch

Image-Based Parking Space Detection and Occupancy Classification [github]
This project proposes a novel way to build a smart parking system that automatically locates valid parking spaces and can monitor and manage parking lots to ensure vehicles are in the right spot. It also classifies regular and accessible parking spaces to improve the accessibility of people with disability to the parking spots. This is implemented by fine-tuning a Mask R-CNN model to extract parking spots as well as classify their occupancy. The Action-Camera Parking dataset is used as the base dataset for parking spot detection and classification and has been increased to improve the model performance. A new custom dataset is also proposed to detect and localize painted handicap parking marks on the ground.
Deep Learning, Fine-tuning, Mask-RCNN, Faster-RCNN, PyTorch, ACPDS dataset, COCO Evaluator

Few-shot Personalized Image Editing using Stable Diffusion (Instruct2Dream) [github]
Developed a Few-shot Personalized Image Editing model utilizing Stable Diffusion to enable precise image modifications based on textual descriptions and reference images. This model uniquely learns and identifies a specific subject through a few samples, allowing for customized edits that closely match user intentions. Enhanced by a custom dataset with input-output pairs and text instructions, our approach finely tunes diffusion models for personalized, high-fidelity image editing.
Image Editing, GenAI, Conditional Generation, Diffusion, Fine-tuning, Few-shot Learning, Custom generated dataset

Spatial Video Grounding via Graph Transformers [github]
Developed a Video Question and Answering (VQA) model leveraging graph transformers, designed to process human queries by pinpointing and emphasizing spatial locations relevant to both the textual prompt and its corresponding answer. This model utilizes a transformer-based architecture to effectively capture video dynamics over time, while the integration of graph neural networks enables comprehensive reasoning across the video to accurately address queries. Furthermore, the model's capability to visually represent information as output enhances evidence-based interpretation, advancing the transparency and reliability of machine learning models.
Deep Learning, Multi-modal (Vision-language), Graph Transformer, Contrastive Learning, PyTorch, Fine-tuning, Localization

Enhancing Urban Mobility via AR Location-based Parking Navigation System [github]
This work presents a novel approach combining an automatic camera-based parking space detector with an Augmented Reality (AR) routing tool to enhance user experience, resulting in an interactive driving navigation system. The proposed solution combines Machine Learning methods for parking space detection with advanced navigation systems using landmarks and pathways, offering benefits beyond time-saving for drivers.
Augmented Reality, ARCore, Geospatial API, GPS, Visual Positioning Sys. (VPS), Camera Calibration, Unity, C#, PyTorch, Flask, AWS

CERTIFICATES

PyTorch Ultimate 2023 *Aug 2023*
Certified by Udemy

Design and Architect Python Apps *Feb 2022*
Certified by Udemy

Python for Computer Vision with OpenCV and Deep Learning *Jan 2022*
Certified by Udemy

Machine Learning Foundations: A Case Study Approach *May 2021*
Certified by Coursera

Deep Learning Fundamentals Course *Jul 2019*
Certified by ACM Student Chapter

Image Processing in Robotics Course *Jun 2017*
Certified by KN2C Robotic Team

SKILLS

Technical Knowledge
Machine Learning, Deep Learning, Transfer Learning, Multi-task Learning, Computer Vision, Large Language Models (LLM), Knowledge Distillation, Video and Image Processing, Parallel and Distributed Training, Stable

Programming Languages
Python (PyTorch, CUDA, CuDNN, TensorFlow, Keras, NumPy, SciPy, PyQt, OpenCV, VTK, ITK/SITK, Matplotlib), C/C++, Java, C#, SQL, HTML, CSS, JavaScript (jQuery), PHP (Laravel), Ruby (Rails)

Diffusion (Generative AI), Real-time Detection, Tracking, Perception, Classification, Segmentation, Software Engineering, Shell Script, Version Control System (Git)

Development Environments

PyCharm, VS Code, IntelliJ Idea, NetBeans, Unity, Google Colab, Jupyter Notebook, Docker, macOS, Linux, Windows

Soft Skills

Teamwork, Cross-functional Collaboration, Communication, Presentation, Deliver Results

REFERENCES

Will be provided upon request.