Qiujie Cui

qjcuis@gmail.com, Seattle, WA, USA

Objective

Machine Learning Engineer with over 10 years of experience in Machine Learning, Robotics, and Computer Architecture, seeking full-time or part-time research opportunities to further innovation in these domains.

Work Experience

Machine Learning Engineer Amazon, Seattle, WA, Sept, 2021 - Present

- Developed a CatBoost-based ML model to rank robot manipulator grasp points using 50+ features from real production, trained on 400K clustered package scenes with AutoGluon. Improved package grasp success rate by 1.5% on 5M packages handled daily, optimized inference speed by 2.5x (0.2ms per grasp) with C++, and established a data curation pipeline for ongoing retraining to mitigate data drift.
- Designed and implemented annotation management service as a building block to Amazon Robotics ML pipeline that dynamically adjusts annotation workloads based on workforce capacity, and enables importance-based data sampling on Amazon S3 via DynamoDB.
- Led a software team in designing an orchestration-based item buffer system for automated item acceptance, storage, reorientation and vending, enabling integration with the Vulcan stow robotic workcell, achieving a throughput of over 300 units per hour..
- Lead for Vulcan stow alpha model integration with Fulfillment Center ecosystem, drove efforts across teams and functions to assess needs, plan implementation and validation, while minimizing overhead cost.

Software Engineer Intel, Austin, TX, July, 2013 - Oct, 2020

- Developed software models to simulate and validate architectural components of the Atom CPU line. Contributed to the modeling of 5 generations of high-performance, low-power CPU architectures, collaborating with design and architecture teams on functional and performance enhancements.
- Responsible for architecture modeling & simulation, feature A /B testing, functional validation, developing test methodologies, performing coverage analysis, bug reproduction, model updates, and maintaining test environments.

Academic Experience

Georgia Institute of Technology Sept, 2017 - June, 2021

- M.S in Computer Science - Perception and Robotics (Major), Machine Learning (Minor), GPA: 4.0

University of Michigan September, 2011 - May, 2013

- M.S in Electrical Engineering – Computer Architecture & VLSI (Major), Embedded Systems (Minor), GPA: 3.7

Latest Projects

Quipster Chatbot

- Developed an LLM-based chatbot using Amazon Q and Retrieval-Augmented Generation (RAG) to efficiently index and search internal wikis and Quip documents through fast keyword-based queries.

Robot Arm Pick and Place

- Developed ROS service nodes for Forward and Inverse Kinematics solvers for the KR-210 (6 DOF) robot, facilitating item retrieval from random shelf locations and placement in designated bins. Utilized Gazebo simulation and MoveIt for route planning and collision avoidance.

3D Perception – Object Detection Using Point Cloud Data and Filter Pipelines

- Implemented a filtering pipeline (statistical outlier filter, down-sampling, passthrough filter, and RANSAC plane fitting) on Point Cloud data from a ROS camera node capturing a simulated desktop in Gazebo.
- Utilized KD-tree clustering to group filtered points and trained a Support Vector Machine (SVM) using HSV histogram and surface vector features to classify each Point Cloud cluster.

Latest Projects (cont'd)

Follow Me – Human Target Detection and Tracking via Semantic Segmentation

- Collected training and validation data with drone filming person of interest walking in a crowd with different backgrounds in the Unity simulation environment
- Designed and trained neural network (depth of 4 encoder/decoder layers and 1x1 convolution layer) for semantic segmentation – to parse a test image and identify the target, crowd and background in the image pixels for drone to follow the target.

Mountain Rover – Terrain Mapping and Mineral Exploration

- Collected image data of a rover treading in a territory with mountains, sand and rock samples in the Unity simulated environment.
- Implemented color threshold and coordinate transformation algorithms to update area map (bird's eye view) with open path, obstacles and rock samples.
- Implemented algorithm to drive rover in direction of open path.

Video target tracking using Kalman filter and Particle Filter

- Tracked target person/object in video footage using continuous template update as target moves in/out, turns or even being occluded for a short period of time.
- Successful tracking using both Kalman filter and Particle Filter modeling of target velocity with measurement noise and movement uncertainty

Video Stabilization using L1 trend fitting

- Extracted frames from shaky footage and computed original camera path, and applied L1 trend fitting using 1st order, 2nd order and 3rd order derivative to generate desired smooth path with frame constraints.
- Back-calculated corrective matrix for each frame and applied to original footage to get stabilized video

Fashion Recognition with Deep Neural Network

- Trained a deep neural network over Deep Fashion Data set (800K labeled images) to recognize and categorize fashion items to a top3 accuracy of 86% among 50 categories
- Based off ResNet-50, split the 2nd-to-last layer into bounding-box prediction and category prediction training to improve accuracy; break data set into 4 groups to train and cross-validation

Warehouse Robot Navigation and Object Retrieval Simulation

- Simulated a robot with noisy sensor and actuator to move around in an unknown warehouse with a given number of boxes to pick up
- Path planning using A* search algorithm and smoothing. Using online SLAM to map environment and localize robot

Isolation Game AI Player

- Implemented game bot for a two-player isolation game using MiniMax strategy to maximize evaluation function outcome for game bot
- Optimized game bot via decision-tree alpha-beta pruning, iterative deepening, and custom evaluation function to beat naive bot 99.9% of time

Flying Spy Shark – Camera-Equipped Toddler Safety Tracker

- Developed and installed a spy camera on an air-swimmer Shark balloon to stream live video via Wi-Fi to a computer, while tracking a toddler target. Utilized Intel Edison embedded platform to execute OpenCV-based video processing and control DC motors for the shark's fin and tail movements.

ASIC Design of Ultra-Low Power FIR Filter with Custom SRAM Design

- Developed a 16-channel digital FIR filter bank, progressing from RTL design to circuit implementation, with optimized data paths for sub-threshold operation. Designed a custom sub-threshold 8-T SRAM incorporating bit-line boosting and body biasing techniques. Implemented dynamic voltage scaling with PVT variation tolerance for SRAM using Canary Logic.

Mentorship

Intel Education Service, USAID-JCP Amman, Jordan, May, 2015

- Led a 2-week workshop on Intel Galileo Gen2 embedded platform, with 70+ participants from 12 universities across Jordan
- Developed training material and conducted Galileo-based lab sessions to show how to work with various types of sensors and actuators, GPIO driver libraries and Linux programming mode
- Float-around debug with different teams to work towards their final projects

Awards

Intel Developer Forum Creative Design Award Austin, Texas, Jul, 2015

- For Flying Spy Shark among other project demos from Intel Austin Makers

AMD Design contest 1st Prize Ann Arbor, Michigan, Apr, 2013

- For ASIC design of an ultra-low-power Sub-threshold variation-tolerant 110-tap FIR filter with custom-designed SRAM

Skills

Languages: Python(numpy/scipy/openCV/sympy/cvxpy), Java, C/C++, Kotlin, System Verilog

ML techniques: LLM, RAG, NLP, Deep Neural Networks (Mask R-CNN/ResNet,etc), Reinforcement Learning, Decision-trees, KNN, K-means, SVM

ML tooling: Langchain, AWS Sagemaker, Amazon Q, Amazon Bedrock, Hugging Face, Pytorch, Keras, Caffe, TensorFlow

Robotics Systems: Carbon(Amazon Internal), ROS

Simulation Environment: Gazebo, Rviz, Unity, MoveIt!