Yiqian Li

Email: 120090702@link.cuhk.edu.cn | Mobile: (+86) 13660576272

EDUCATION BACKGROUND

The Chinese University of Hong Kong (CUHK), China 09/2020 – 07/2024
BS Computer Science and Engineering | Overall GPA: 3.6/4.0

The University of California, Irvine (UCI), United State 01/2023 – 06/2023
BS Computer Science and Engineering | Overall GPA: 4.0/4.0

Core Coursework: Multivariable Calculus | Probability and Statistics | Linear Algebra | Optimization | Discrete Mathematics | Data Structures | Design and Analysis of Algorithms | Graph Models | Operating System | Database System | Machine Learning | Intro to AI | Project in AI | Formal Language and Automata | Computer Networks | Microprocessors and Computer Systems |

SKILLS

Programming Languages: C/C++ | Java | Python | MATLAB | SQL | ARM Assembly | AHDL | Verilog HDL
Libraries: NumPy | SciPy | Pandas | PyTorch | Scikit-Learn | TensorFlow | Matplotlib | Seaborn | Selenium | libtorch
Software: MS Office | LaTex | CUDA | PyCharm | Eclipse | IntelliJ | Git | Tableau
Membership: Student Robot Association | UAV and Intelligent Vehicle Society | Computer Society

COURSE PROJECTS

Operating System **Course Project, CUHK**

- Gained hands-on experiences in virtual memory management, virtualization, concurrency, synchronization, file systems, heterogeneity and security:
  - Implemented a mechanism of file system management via GPU's memory
  - Simulated a mechanism of virtual memory via GPU's memory and CUDA API
  - Built a disk request scheduler based on mutex and conditional variables for handling requests from multiple threads
  - Built a prime device in Linux, and implemented file operations in kernel module to control the device

Database System **Course Project, CUHK**

- Built a database system for real-time assignment of chip manufacturing orders to plants varying in production capacity:
  - Designed database structure using ER diagram; designed an order distribution algorithm, allowing real-time update; built web pages to display public production information
  - Gained familiarity with the design, implementation, testing, and maintenance of relational database system, focusing on efficiency, security, and robustness

Introduction to AI **Course Project, UCI**

- Implemented various approaches in Python to solve Sudoku as a constraint satisfaction problem based on five algorithms, including variable selection heuristics, value selection heuristics, consistency checks, and extra heuristics

Introduction to Graphical Models **Course Project, UCI**

- Designed, implemented and characterized probabilistic models for estimating the skill of a game player
  - Enabled the modeling of the player's skill based on continuous (Gaussian random variable) and discrete variables (using pyGM)
  - Benchmarked the model's predictive power against those of simple approaches; ran sensitivity of game number on predictive accuracy; further enhanced model's complexity by accounting for game features and additional latent scores

Reinforcement Learning for Optimal Traffic Signal Control
Project in AI **Course Project, UCI**

- Designed, implemented, and validated in Python a set of reinforcement algorithms for optimal traffic signal control so as to minimize waiting time of vehicles
  - Simulated cross-road traffic condition using Simulation of Urban Mobility (SUMO) platform with built-in maps, allowing vehicle flow visualization
  - Employed 1) Deep Q-network, Double Deep Q-network, Dueling Deep Q-network, and Proximal Policy Optimization (PPO) reinforcement learning algorithms, and 2) various exploration strategies including greedy, epsilon-greedy, and upper confidence bound (UCB) to drive control, followed by a comparative study on the effectiveness and efficiency

Machine Learning **Course Project, CUHK**

- Established a data-driven workflow to predict the song genres via supervised learning:
  - Mined 50K+ songs from Spotify, followed by data cleansing to 1) drop data with no causal influence, 2) handle missing data points, 3) categorical data encoding, 4) data scaling via z-score standardization, 5) principal components analysis (PCA) for dimensionality reduction, and 6) data clustering and tSNE for visualizing high-dimensional data.
  - Trained SVM, random forest, and deep neural network models to drive prediction; evaluated model performance using ROC curves and AUC scores.

Computer Networks **Course Project, CUHK**

- Designed and implemented in Python a multi-host real-time file sharing app, with a focus on delivering 1) decent flexibility in handling various file formats, 2) high transmission efficiency and satisfactory credibility, and 3) compatibility with various transmission objects:
  - Enabled real-time data sharing between three virtual machines via LEFT mechanism while delivering supreme data integrity and speed
  - Allowed 1) file sharing via centralized server structure and TCP protocol, 2) breakpoint resume transmission via heartbeat mechanism, 3) file encryption via AES, and 4) file compression.

***Computer Architecture* Course Project, CUHK**

- Collectively leveraged MIPS instruction set, assembly language, Vivado software, and Verilog language to design a single-cycle CPU, a multi-cycle CPU and a 5-stage pipelined CPU, covering control unit, internal registers, ALU, and instruction set; validated the CPU's functionality and performance by running an array of computer programs.

EXTRACURRICULAR ACTIVITIES

Implementation of a Flight Control System Based on Kalman and PID Controller
CUHK UAV and Intelligent Vehicle Society 01/2023 - 03/2023

- Implemented in MATLAB an integrated flight controller based on Kalman filter and PID controller to allow the speed and roll/yaw/pitch angles control of a fixed-wing unmanned air vehicle (UAV) with short transition, decent stability and good anti-disturbance characteristics.
- Key steps include 1) derivation of flight dynamics model and transfer function of UAV, 2) PID controller design, 3) Kalman filter for UAV state estimation, and 4) simulation of PID controller and Kalman estimator using MATLAB Simulink.

Kaggle Competition: SIIM-ISIC Melanoma Classification 01/2021 - 03/2021

- Built a machine learning model for classifying melanoma based on skin lesion images：
  - Conducted image preprocessing using 1) CutMix as a regularization strategy to train strong classifiers with localizable features, and 2) mixup for training a neural network on convex combinations of examples pairs and their labels.
  - Evaluated an array of segmentation models, including 1) U-Net CNN, and 2) Google's DeepLab semantic image segmentation model with deep convolutional nets, atrous convolution, and fully connected CRFs.
  - Validated the model with five-fold cross validation, followed by hyperparameter tuning to improve prediction accuracy

***Team Member* | DJI RoboMaster Robotics Competition** 10/2020 – 05/2021

- Joined a multi-disciplinary team to prepare for the DJI Robomaster competition, developing and operating six robots that launch projectiles at opponent in order to deduct Hit Points:
  - Established a simulation environment in Gazebo simulator to facilitate the intelligent robot development and testing
  - Analyzed QMIX and VDN algorithms for intelligent vehicle decision, covering engineering (loading and transporting bullets), infantry (fight with other robots), supply (ammunition management) and aerial (air combat) robots
  - Enabled route planning via A* algorithm, Dijkstra's algorithm, and a greedy heuristic

RESEARCH EXPERIENCES

Medical Image Synthesis and Denoising based on Diffusion Probabilistic Model
***Research Assistant* | UCI** 06/2023 - Present

- Establish and validate a 2D medical image synthesis framework based on a diffusion model, covering a forward Gaussian noise process and a reverse denoising process using the transformer-based diffusion model.
- Systematically evaluate the quality, diversity, and authenticity of the synthesized medical images based on a set of fit-for-purpose metrics, covering X-ray, MRI and CI images.

Computer-Assisted Analysis of Sperm Morphology
***Research Assistant* | CUHK** 12/2022 - Present

- Applied U-net for sperm segmentation.
- Applied DBSCAN and Agglomerative Clustering for single sperms image sets and single standard normal sperm image to generate pseudo labels.
- Applied Resnet and Vision-transformer as semi-supervised classification models for single sperms image sets with pseudo labels.

Sperm Video Segmentation and Tracking
***Research Assistant* | CUHK** 08/2022 – 12/2022

- Optimized the up-sampling process in U-Net to accelerate segmentation of sperms in each video frame.
- Enhanced the occlusion-flag-enabled PFT model for improved recognition and tracking of sperms in videos by modifying the occlusion-judge mechanism; leveraged TensorRT for model acceleration.
- A research manuscript submitted to IEEE Transactions on Automation Science and Engineering.

Design of a Log Detection and Analysis System for Huawei Cloud Service
***Research Assistant* | CUHK** 09/2021 – 07/2022

- Developed a log detection and analysis system for Huawei cloud service:
  - Implemented a semantic parser to parse logs; applied word2vec to calculate semantic similarity of logs to agglomerate logs from the same root cause.
  - Extracted invariants among logs, followed by monitoring the dependency of invariants to inform log detection

Log Paring Analysis
***Research Assistant* | CUHK** 11/2021 – 05/2022

- Analyzed false detection by log parsing algorithms such as Drain, AEL, and IPLoM; a research paper submitted to International Symposium on Software Reliability Engineering (ISSRE).
- Corrected some ground truth in log parsing datasets; a research paper submitted to the International Conference on Software Engineering (ICSE).
- Evaluated the accuracy and efficiency of CNN, auto decoder and LSTM algorithms for automated anomaly detection of logs.