CHONG HU

(917)-388-5186 | ch3467@columbia.edu
https://jacksnowwolf.github.io/ | https://www.linkedin.com/in/chong-hu-5b1a6514b/
Apt. 4A, 169 Manhattan Ave, New York, NY, 10025

EDUCATION

Columbia University, New York, US *Aug 2019 - Dec 2020 (expected)*
The Fu Foundation School of Engineering and Applied Science
M.S. in Electrical Engineering

Shanghai Jiao Tong University (SJTU), Shanghai, CN *Sep 2015 - Aug 2019*
Joint Institute: University of Michigan-Shanghai Jiao Tong University Joint Institute (UM-SJTU JI)
B.S. in Electrical and Computer Engineering; Minor in Data Science
Courses: Data Structures and Algorithms, Operating System, Methods and Tools for Big Data, AI Techniques

WORK EXPERIENCE

MokaHR Company, Beijing, CN *Dec 2018 - Apr 2019*
Software Engineer Intern, AI Group

- Combined CTPN and CRNN and developed model to solve OCR problems (Chinese & English) in resume images using TensorFlow; simplified Network Structure and sped up inference time 2s/10s on average, lost only 2% accuracy
- Adapted open source labeling software to mark text and run evaluation and unit test for different stages
- Packaged model into web service using gunicorn and Flask, provided API and deployed on Alibaba cloud
- Implemented cache mechanism with Redis and multistage recognition with high accuracy (over 90% per label) Improved 15% overall performance and about 200% QPS over the original third-party service with parallel processing in Python

Beijing Infervision Company, Beijing, CN *Jan 2018 - May 2018*
Software Engineer Intern, Modeling Group

- Applied YOLO V2 & V3 under darknet frame and FPN under MXNet for illness detection on DR images
- Calculated anchor size and number in different methods for YOLO and combined three detection layers to improve accuracy by roughly 5% on tiny objects
- Utilized Focal Loss to replace original softmax function to focus on cases with fewer samples; increase average accuracy by about 3%

PROJECTS

High Dynamical Range (HDR) Video Recovering Algorithm *Aug 2018 - Dec 2018*
Deputy Team Leader, SJTU *Graduation Project*

- Used hdrcnn model to train data to transform LDR to HDR with different data enhancements and loss function (e.g., cosine loss), in order to reconstruct over exposed area and restore details in dark area
- Evaluated model performance using HDR-VDP v2 and obtained 20/100 more than traditional method
- Applied FFmpeg, OpenEXR to finish the transfer from LDR video to image and image to HDR video; added meta data of HDR10 format and corresponding BT2020 curves

Music Recommendation System Analyzed from Million Song Dataset (MSD) *Jun 2019 - Aug 2019*
Team Member, SJTU *Course: Methods and Tools for Big Data*

- Deployed Hadoop with Spark and Drill and extracted song information from 160GB avro files containing h5
- Built similar artist adjacent matrix using MapReduce in Hadoop and Spark, used Naive Bayes to guide the scaling data; ran hierarchical and k-mean++ clustering methods to split the genres of different music
- Visualized results in Matplotlib in Python and ggplot2 in R and constructed music recommendation logic

Multi-threaded and Efficient Programming in Database *Oct 2018 - Dec 2018*
Team Member, SJTU *Course: Introduction to Operating System*

- Implemented table management queries and data manipulation in C++ and handled exceptions in query error
- Accelerated database using multi-threading and optimizing data structure, enabling 50% faster speed

TECHNICAL SKILLS

Programming Language: Python, C++, C, R, Java, MATLAB, Julia, SQL, Verilog.
Toolkits/Frameworks: Linux, Hadoop, Drill, Spark, Git, NumPy, pandas, TensorFlow, Matplotlib, OpenCV