Yunyi Shang

Phone: (+86) 137 9898 8355 | Email: winnyshang@gmail.com
https://winny-shang.github.io/

EDUCATION

Macau University of Science and Technology **Macau, China**
Master of Science in Computer and Information Systems *09/2018 – 08/2020*

- GPA: **3.97**/4.0 | Rank: 1st/35 | Core Courses: [Data Warehouse and Data Mining A+, Software Engineering A, Research Methods in Information Technology A+, Computer Architecture A+]

Beijing Normal University, Zhuhai **Zhuhai, China**
Bachelor of Engineering in Computer Science and Technology *09/2014 – 07/2018*

- GPA: **4.0**/5.0 | Rank: 1st/97 | Core Courses: [Discrete Mathematics 94/100, Data Structure 95/100, Operating System 92/100, Compiling Principle 96/100, Probability and Mathematical Statistics 91/100]

RESEARCH EXPERIENCE

Automatic Portrait Image Pixelization **03/2019 – 08/2020**
Supervised by Prof. Hon-Cheng Wong, Computer Graphics and Vision Group, Macau University of Science and Technology
Objective: Maintaining sharp boundaries in low-resolution while using a limited color range to convey the image content.

- Extracted portrait features using the Cascade Object Detection algorithm.
- Abstractly represented images through non-uniform five-dimensional clustering, considering spatial and color aspects, based on pixel art characteristics. Algorithms used: SLIC Superpixels, FISODATA.
- Enhanced image effects using algorithms such as Gaussian Bilateral Filtering, Sobel Edge Detection and Morphological Algorithms.

Conclusion: By combining the strengths of existing methods, a novel approach is proposed for automatic pixelated image generation. This method incorporates image downscaling techniques and color gradient extraction methods to enhance visual effects and meet the aesthetic requirements of pixel art.

Improved Viola-Jones face detection algorithm under the HoloLens headset **03/2017 – 03/2018**
Supervised by Prof. Jing Huang, BNU High Tech, Beijing Normal University, Zhuhai
Objective: Hololens Device-Based Facial Detection for Real-Time Applications.

- Boosted the detection speed of the classic Viola-Jones face detection algorithm using two-dimensional convolution separation.
- Optimized detection speed in the application by leveraging image resampling techniques with Microsoft's mixed reality headset, HoloLens.
- Leveraged SLAM (Simultaneous Localization and Mapping) technology for 3D positioning, enabling better human-computer interaction effects through gesture interaction.

Conclusion: Compared to existing face detection interfaces, the accuracy has increased by 12%, and the detection speed has improved fourfold.

PUBLICATIONS

- **Shang Y**[1], Wong H C. Automatic portrait image pixelization[J]. Computers & Graphics, 2021, 95: 47-59.
- Huang J, **Shang Y**, Chen H. Improved Viola-Jones face detection algorithm based on HoloLens[J]. EURASIP Journal on Image and Video Processing, 2019, 2019(1): 1-11.

WORK EXPERIENCE

Tencent Technology (Shenzhen) Co., Ltd **Shenzhen, China**
Applied Researcher *09/2020 – 08/2022*

Project: Intelligent Classification of Feedback Text
Objective: The project employs a cascaded approach, combining machine learning and deep learning models, to achieve multi-level classification of feedback text. It involves algorithm development and optimization from scratch to completion.

- Employ the Term Frequency-Inverse Document Frequency (TF-IDF) algorithm to extract keywords from historical user feedback text. Calculate text similarity using hierarchical clustering to generate a hierarchical nested clustering tree.
- Developed a novel model structure combining Convolutional Neural Networks (CNN) and Transformer models. Leverage BERT as the foundation and add convolutional layers for multi-level feature extraction. Employ a hierarchical approach that combines convolution and global attention for local and global feature extraction, creating a text-based classification model.
- In conjunction with the TF-IDF probability model, leverage text keywords to predict the classification of tail classes from a probabilistic perspective. This improves the recognition rate of tail categories while maintaining high accuracy for head classes in the neural network model and reduces the imbalanced categories problem.

Conclusion: After deployment, the project saved more than 50% reduction in the operational workforce. The algorithm achieved an average accuracy of 85% and improved the ability to identify user feedback issues by 30%. The response time for addressing user queries was reduced by 90%.

Project: Log Anomaly Detection

Objective: Combine online analysis capabilities and anomaly detection algorithms to provide a generalized solution for problem analysis, prediction, and recommendation. This improves problem localization efficiency for businesses.

- Collect historical terminal logs and extract log content templates using the LogSig algorithm, compressing a large volume of logs with a pre-defined template. Utilize PCA calculating Squared Prediction Error to detect anomalies in individual log text, enabling unsupervised anomaly mining.
- Design a text feature extractor based on log sentiment lexicons and anomaly information. Conduct sentiment mining on log texts and train a binary classification model based on BERT for log sentiment, achieving an accuracy of 90% on standard datasets.
- Employ an asynchronous model inference approach by invoking the model offline on servers. Optimized the model inference speed by using TensorRT to lower the latency. Asynchronous inference alleviates server and model pressure, and store inference results in a database for later viewing by developers.

Conclusion: After deployment, the project handled more than 5,000 log files per day in average, detecting more than 47,000 lines of anomalous logs. Optimizations were made to log downloading, decompression, parsing, searching, and other capabilities, reducing user time costs.

Project: High-Frequency Feedback Alert System Reconstruction

Objective: Reconstruct and upgrade an outdated project to monitor newly emerging hotspots in user feedback in real-time by employing condition filtering, threshold judgment, fluctuation detection, and algorithm mining of high-frequency words.

- Streamline project functional modules and extract their interaction relationships, transforming procedural-oriented design into object-oriented design.
- Based on the Domain-Driven Design (DDD) architecture, restructure the project into different domain modules such as feedback management, alert monitoring, notification, etc. Rationally divided and decoupled services.
- Utilize the trpc-go development framework for backend microservice reconstruction. Migrate the original project's self-built databases (MongoDB, Redis, etc.) to cloud and optimize operational resources.
- Employ the trpc-python development framework for microservice reconstruction of alert algorithm scripts. After reconstruction, use a message queue (MQ) to implement strategies such as peak smoothing, concurrency, monitoring, and failure retry.

Conclusion: The reconstructed services demonstrate improved stability and response speed. The new services expand alert configuration capabilities by providing multi-dimensional combinations of alert conditions, allowing for more flexible alert configurations.

INTERNSHIP EXPERIENCE

Tencent Technology (Shenzhen) Co., Ltd **Shenzhen, China**

Data Analysis and Backend Development *06/2020 – 09/2020*

Task 1: Feedback Information Preprocessing

- Employ TF-IDF algorithm to feedback texts and extract image features using the localHash algorithm, which are then stored in the database. Filtered similar contents based on cosine similarity of texts and hamming distance of images between the user query and all of the feedback from the past three months. Introduce real-time calculation of spatial vectors for new feedback, comparing them with feedback from the past three months.

Conclusion: Achieved a 65% reduction in repetitive issues, increased duplicate detection rate by 122%, and optimized the input-output ratio of feedback problems.

Task 2: Data Statistics

- Developed a scheduled service to retrieve historical feedback and associated issue tickets, query the real-time status of issue resolution, and update the database accordingly.
- Perform weekly scheduled database queries and generate reports summarizing the handling status and proportions of user feedback for the week. Automated distribute these reports to relevant personnel via email.

Conclusion: Transitioned from 2 person-days of offline statistics to real-time statistics, improving accuracy and generating traceable source data.

AWARDS AND HONORS

- Graduate Scholarship (50% tuition fee waiver), (awarded by Macau University of Science and Technology, September 2018 - September 2020)

- Beijing Normal University, Zhuhai Outstanding Academic Scholarship, Ranking: 1.50% (November 2015, November 2016, November 2017)
- Outstanding Graduate, Outstanding Graduation Thesis (awarded by Beijing Normal University, Zhuhai, June 2018)
- Third Prize in the Guangdong Provincial College Computer Works Competition for the Improved Viola-Jones Face Detection Algorithm based on HoloLens (awarded by the Guangdong Computer Society, June 2018)
- ACM Programming Contest Excellence Award, IT Festival Second Prize (awarded by Beijing Normal University, Zhuhai, 2016, 2018)
- Excellent Community Work Award (awarded by Beijing Normal University, Zhuhai, November 2015, November 2016, November 2017)
- Outstanding Peer Mentor (awarded by Beijing Normal University, Zhuhai, October 2016)

PATENTS

- **Yunyi Shang**, Ruolin Zheng, Zongxing Xie, Jiayi Wei. A feedback information classification method, apparatus, electronic device, and storage medium. CN2022103576915
- **Yunyi Shang**, Ruolin Zheng, Zongxing Xie, Guang Song, Jiayi Wei. Training Method, Detection Method, and Apparatus for a Log Detection Model. CN2022115234904

EXTRACURRICULAR ACTIVITIES

Computer Graphics **09/2019 - 01/2020**

Teaching Assistant, Macau University of Science and Technology

- Providing on-site guidance for experimental operations.
- Grading assignments and providing explanations for questions.

BNU High Tech Team / BNUZ VR Experience Store **03/2017 – 03/2018**

Team Leader / Store Manager, Beijing Normal University, Zhuhai

- Coordinated and communicated with team members, organized and supervised work plans.
- Led several projects, including a national-level AR map navigation system based on iOS and a commercial AR word recognition card based on the Android system, among others.

School of Information Technology Students' Union **06/2015 – 06/2016**

Director of Public Relations, Beijing Normal University, Zhuhai

- Successfully pursued and secured sponsorships for various events and activities, obtaining the necessary financial support and resources.
- Managed the official public WeChat account of the Students' Union and published content.
- Organized the Student Representative Assembly and oversaw the logistical aspects.

SKILLS

Data Analysis:

- Programming Languages: C, C++, C#, Python, Java, Golang, Matlab
- Machine Learning/Deep learning: PyTorch, TensorFlow, Scikit-Learn
- Framework: RPC, MVC, DDD
- Database: MySQL, Redis, MongoDB
- Others: Android Studio, Unity

Language Proficiency:

- English (TOEFL: 101, GRE: 326 - V153, Q169, AW3)
- Mandarin and Cantonese (Native Speaker)