Shuzhao Xie

(+86) 188-1025-9895 • xiesz99@gmail.com • shuzhaoxie.github.io

Education

Tsinghua University Shenzhen, China
M.E. in Computer Technology 08/2020 – 06/2023

Beijing Normal University Beijing, China
B.S. in Computer Science and Technology, 09/2016 – 07/2020

Research Interest

My interest lies in designing multimedia content synthesis, compression, and transmission algorithms for real-time interactive applications. Specifically, I am interested in NeRF-based 3D synthesis and 2D/3D/volumetric video streaming. My other interests include distributed inference and respiratory sound classification.

Experience

Multiple MLaaS Inference Shenzhen, China
Advisor: Zhi Wang, Yifei Zhu 07/2020 – 10/2021

- Measured the performance (e.g., accuracy, throughput, and latency) of machine learning inference services (mainly object detection services) on AWS, Azure, Google Cloud, and Alibaba Cloud.
- Found the higher cost but occasionally lower accuracy problem under the multi-cloud environment and solved this NP-hard problem with combinatorial reinforcement learning algorithms.
- Evaluation results show that our solution decreases the cost by 67% without any sacrifice to accuracy.

Deep Learning Model Acceleration Shanghai, China
Advisor: Chenyang Guo 01/2022 – 06/2022

- Analyzed user data from Hupu and appended TFRecord to reduce the latency of model training.
- Implemented Faster-RCNN and converted it into TorchScript with LibTorch and CUDA.
- Tested the latency and accuracy of object detection models on PyTorch, ONNX, OpenVino, and TensorRT.

Respiratory Sound Classification Shenzhen, China
Advisor: Zhi Wang, Zhiyong Wu 06/2022 – Now

- Extracted acoustic features, such as Mel Spectrogram, MFCC, and Energy Summation.
- Implemented a series of models and papers, including Glance-and-Gaze, Faster-RCNN, and Conformer.
- Deployed abnormal sound detection models into the Alibaba Cloud and private cloud clusters.
- Led the technical team and coordinated the private dataset collection process.

Publications/Preprints

- **Shuzhao Xie**, Yuan Xue, Yifei Zhu, and Zhi Wang. "Cost Effective MLaaS Federation: A Combinatorial Reinforcement Learning Approach", **INFOCOM'22**, [Page] [PDF] [Code]
- **Shuzhao Xie**, Yuan Xue, Yifei Zhu, and Zhi Wang. "SkyML: An Inter-Cloud Broker for MLaaS Federation", [In Submission to IEEE Transactions on Multimedia]

Teaching Experience

Big Data System (B) Tsinghua University, SIGS, China
Teaching Assistant, Advisor: Zhi Wang 09/2021 – 01/2022

- Implemented a distributed file system (DFS) with light fault tolerance.
- Implemented the MapReduce framework based on the above DFS.
- Implemented a linear regression algorithm with Spark.

Selected Awards

Bronze Prize, ACM-ICPC Asia Regional Contest (Jiaozuo Station) 2018
Bronze Prize, CCPC Regional Contest (Qinhuangdao Station) 2018

Skills

Programming Languages: Python, C/C++, Bash, XML, Java
Tools and Frameworks: Git, LaTeX, PyTorch, CUDA, Docker, Spark, Android