

RESEARCH INTERESTS

Video Analytics, Computer Networking

EDUCATIONS

2019 Sep - present

Ph.D. in Computer Science

UNIVERSITY OF CHICAGO, USA

2015 Sep - 2019 Jul

Bachelor in Computer Intelligence Science

PEKING UNIVERSITY, CHINA

PUBLICATIONS

MLSys 2022

AccMPEG: Optimize Video Encoding for Accurate Video Analytics

Kuntai Du, Qizheng Zhang, Anton Arapin, Haodong Wang, Zhengxu Xia, Junchen Jiang

[Github link](#)

SIGCOMM 2020

Server-Driven Video Streaming for Deep Learning Inference

Kuntai Du*, Ahsan Pervaiz*, Xin Yuan, Aakanksha Chowdhery, Qizheng Zhang, Henry Hoffmann, Junchen Jiang

[Github link](#)

MobiCom 2020

Renovating Road Signs for Infrastructure-to-Vehicle Networking: A Visible Light Backscatter Communication and Networking Approach

Purui Wang, Lilei Feng, Guojun Chen, Chenren Xu, Yue Wu, Kenuo Xu, Guobin Shen, Kuntai Du, Gang Huang, Xuanzhe Liu

HotMobile 2022

Understanding the Potential of Server-Driven Edge Video Analytics

Qizheng Zhang, Kuntai Du, Neil Agarwal, Ravi Netravali, Junchen Jiang

[Github link](#)

WORK EXPERIENCE

2022 May - 2022 Jul RESEARCH INTERN @ MICROSOFT RESEARCH, USA

Earth+: multi-band satellite imagery compression leveraging historical earth observations

- Measure the potential of only encoding the areas that differ from historical images: $16\times$ compression on over 200 cloud-free images, each covers a $100\text{km}\times 100\text{km}$ area.
- Implement Earth+, a satellite imagery compressor based on historical images. Note that we can compress images with different illumination condition and cloud coverage conditions.
- **Experimental result:** Earth+ achieves $5\times$ compression on one month data obtained from 12 different $100\text{km}\times 100\text{km}$ areas.
- Make a short demo video ([link](#)) that demonstrates Earth+ step-by-step.

2021 Jul - 2021 Sep RESEARCH ENGINEER @ TUSIMPLE, USA

Improving object detection on one video frame by using surrounding video frames

- Perform ablation study based on [state-of-the-art video object detector](#). The performance gain comes from aggregating the DNN features from both nearby video frames and distant (e.g. several seconds ago) video frames.
- Perform parameter sweeping to optimize the object detection performance (measured by mean average precision).
- Profile the whole model and optimize the most time-consuming module (aggregation module) to speed up the training time.
- **Experimental result:** mean average precision +2%, training time -6 hours.

RESEARCH EXPERIENCE

2021 Oct - Present UNIVERSITY OF CHICAGO, USA

Advised by Prof. Junchen Jiang; Tool: **PyTorch**

OneAdapt: Fast Configuration Adaptation for Deep Learning Applications via Backpropagation

- Leverage backpropagation to identify which part of the video data are crucial to inference results.
- Optimize the knobs (e.g. video resolution) while maintaining high fidelity on the crucial data.

- **Experiment result:** reduce the end-to-end delay, including the data encoding, data streaming and data inference delay, by 10-40% without hurting inference accuracy.

2020 Jul - 2021 Oct UNIVERSITY OF CHICAGO, USA

Advised by Prof. Junchen Jiang; Tool: *PyTorch, OpenVINO*

AccMPEG: Optimizing Video Encoding for Accurate Video Analytics

- Train an ultra-cheap neural network to identify areas in the video that are important to video analytic results and use this neural network to compress the video.
- Accelerate this neural network by 20x on CPU through hardware acceleration (openVINO). After acceleration, the runtime of this cheap neural network is only 50 ms per video frame on CPU.
- **Experiment result:** reduce the video streaming delay by upto 40% while maintaining high inference accuracy.

2019 Sep - 2020 Jul UNIVERSITY OF CHICAGO, USA

Advised by Prof. Junchen Jiang; Tool: *Tensorflow, PyTorch*

Server-Driven Video Streaming for Deep Learning Inference

- Leverage the intermediate output of the DNN to identify regions in the video that may potentially contain objects.
- Let the server re-transmit these regions in higher quality to improve the overall inference results.
- **Experiment result:** reduce the bandwidth consumption of video streaming by 59% with higher accuracy.

2018 Sep - 2019 Jul PEKING UNIVERSITY, CHINA

Research Assistant of Prof. Chenren Xu; Tool: *C++, MCU*

Battery-free Backscatter Communication and Localization via Visible Light

- Revised previous demodulation algorithm to support high-frequency demodulation.
- Proposed a localization algorithm based on the demodulation algorithm.
- **Experiment result:** enable demodulation at 1kHz and achieve accurate localization (mean absolute error (MAE) <1m).

2018 Jul - 2018 Aug

UNIVERSITY OF EDINBURGH, UK

Research Assistant of Prof. Boris Grot; Tool: *Java, C++, Boost, OpenMP*

Software Performance Optimization

- Profile the parallel bottleneck of the software: load imbalance between CPUs and too much remote memory access.
- Re-implemented the bottleneck part of the software in C++ for fine-grained workload and memory control.
- Proposed a heuristic algorithm to optimize the memory access patterns.

2017 Jul - 2018 Jul

PEKING UNIVERSITY, CHINA

Research Assistant of Prof. Jiaying Liu; Tool: *Python, Tensorflow*

Image Enhancement under Poor Weather: Image De-raining

- Proposed a rain discriminator to discriminate between rain pattern and object texture.
- Proposed a multi-granularity perceptual loss to examine the image similarity on multiple scales.
- **Experimental result:** enhances the PSNR by 0.6 and SSIM by 0.8.

SELECTED AWARDS

Nomination on Google Ph.D. Fellowship (2020)

Merit Student (2016-2017)

Kwang-Hua Scholarship (2016-2017)

Bronze medal in National Olympiad in Informatics (2014)