Dayou Du

+1 (646) 206-7968| dayoudu@nyu.edu |25 Park Lane South, Jersey City, NJ 07310

EDUCATION

New York University – New York, NY *Sep. 2017 – May 2019 (Expected)*

Master of Science in Computer Science, Courant Institute of Mathematical Sciences

- GPA: **4.0/4.0**
- Courses: Operating Systems, Distributed Systems, Deep Learning, Computer Vision, Real-time and Big Data Analytics
- Tutor in graduate course Fundamental Algorithms

Peking University – Beijing, China *Sep.2013 – July 2017*

Bachelor of Science in Computer Science, Double Major in Economics

- Major GPA: **3.71/4.0**
- Core courses: *Algorithm Design and Analysis, Data Structure and Algorithm, Operating Systems, Computer Organization, Computer Architectures, *Computer Networks, *Software Engineering, *Database Systems, Parallel Computing, Linear Algebra, Set Theory and Graph Theory, Probability Theory and Mathematical Statistics; (*: Honor Track)

SKILLS

Programming: C++, C, CUDA, Scala, Java, Spark, Hadoop, Hive, Impala, Python, OpenMP, MPI, SQL, Html, JavaScript
Tools and Skills: Vim, Visual Studio, IntelliJ, Eclipse; Git, Gprof, Valgrind, Perf; MySQL, MongoDB; Caffe; Working on Linux

INTERNSHIP EXPERIENCE

Megvii Technology Limited. (aka. Face ++) – Beijing, China *Feb. 2017 – May 2017*

Research Development Intern, Engine Group, RSDE

- Optimized CNN primitives and general tensor manipulation operations for a self-developed deep learning framework on various platforms, including x86_64/ARM/CUDA. The framework is the infrastructure of company-wide developments and researches.
- Obtained several times speed-up for convolution and matrix multiplication operations comparing to OpenBLAS library.
- Developed an automatic profiling tool to find the bottleneck operations when running a DNN model.

School of Informatics, CArD group, University of Edinburgh – Edinburgh, UK *Summer 2016*

Research Intern

- Developed a high-level expression generator in Scala to investigate different combinations of high-level functional patterns.
- Testing and debugging the framework using the expression generator, reduced 80% of the crashes/unexpected results.

Institute of High Performance Computing (IHPC), A*STAR – Singapore *Summer 2015*

Research Assistant

- Deployed Caffe and object detection application on mobile platforms equipping mobile GPU cores in C++/CUDA.
- Discovered and investigated the sparsity in the applications. Designed a sparse-dense matrix-mult algorithm based on novel compression format, showed 1.82X speed-up and 46% energy savings compared to the baseline cuBLAS implementation.
- Concluded our work and submitted a paper to ACM TECS, resulted in publication.

Center for Energy-efficient Computing and Applications, Peking University – Beijing, China *Sep. 2014 – May 2017*

Undergraduate Research Intern

- Proposed and implemented an algorithm to improve parallel scalability by eliminating data-transmission latency inter/intra different nodes using C/MPI/Pthreads, which achieved nearly linear scalability on a 12-node TK-1 cluster.
- Modified and Deployed Caffe library on FPGA platform. Extracted and rebuilt CNN-based stereo matching algorithm on FPGA platform, reduced 90% of the memory required by tiling convolution layers.

SELECTED COURSE PROJECTS

Real-time and Big Data Analytics *Fall 2017*

- Using Hadoop to perform scalable sentiment analysis/opinion mining on Twitter data-stream in a real-time fashion.
- Developed a neighborhood quality analysis model based on the collected data using Hadoop Hive, Impala and Spark MLlib.

Database Systems (Honor Track) *Spring 2016*

- Implemented a local community detection algorithm with free rider effect elimination introduced in a frontier paper.
- Improved the asymptotic time complexity by designing a novel algorithm based on the same query-biased points weighting idea.

Software Engineering (Honor Track) *Spring 2016*

- Developed an on-line course resources sharing platform as chief architect in the team.
- Ranked 5th place among 600 teams in “The Way to Silicon Valley” Innovation and Entrepreneurship Competition.

SELECTED PUBLICATIONS

Dayou Du, Xinfeng Xie, Qian Li, etc., Exploiting Sparsity to Accelerate Fully Connected Layers of CNN-based Applications on Mobile SoCs*, ACM Transactions on Embedded Computing Systems(TECS), Volume 17 Issue 2, January 2018, Article No.37*

GRANTS/HONORS/AWARDS

- Ranked 8th in the 3rd ProgNova Programming Contest (ICPC division) 2017
- LeetCode Weekly Contests Rating - Global Ranking top 1% 2017
- Guanghua Scholarship (5%) ; PKU Research Excellent Award (5%) 2016
- The 3rd Prize, “The Way to Silicon Valley” Innovation and Entrepreneurship Competition (top 5 out of 600 teams) 2016
- Honorable Mention, Mathematical Contest in Modeling 2015
- Grants of National Students Innovation and Entrepreneurship Training Program 2015
- Recognition Award, PAC National Parallel Application Challenge 2014
- The 3rd Prize, 13th "Schlumberger Cup" ACM Programming Contest 2014
- PKU Social Practice Excellent Award (5%) 2014