

Bitan Hou | Curriculum Vitae

7F012, Bldg #25, 328 Xinghu St., Suzhou Industrial Park, P.R. China

🌐 www.houbitan.com

✉ Bitan.Hou@microsoft.com

☎ (+86) 182-1752-7101

Profile

Microsoft STC Asia

Full Time Employee, Software Engineer, Report to Serena Chen

Suzhou, China

Sep 2020 – Present

Deep Motion

Full Time Employee, Deep Learning R&D Engineer

Beijing, China

Nov 2018 – Aug 2020

Microsoft Research Asia

Intern, System Research Group

Beijing, China

July 2018 – Nov 2018

Shanghai Jiao Tong University (SJTU)

Bachelor of Engineering, Outstanding Graduate (Top 3%)

Shanghai, China

Sep 2014 – July 2018

- School: Electronic Information and Electrical Engineering, Department: Electronic Engineering (EE)
- GPA of Upper Division Work: 3.83/4.3(89.59/100), Standing: 4/60

Publication

- **Bitan Hou**, Yujing Wang, Ming Zeng, Shan Jiang, Ole J. MengShoel, Yunhai Tong, Jing Bai.
Customized Graph Embedding: Tailoring Embedding Vectors to different Applications. [arXiv]

Work&Research Experience

- **Whole Page Optimization**
 - Working on the Microsoft Bing search engine, MSN, and New tab of Edge browser to optimize the user experience, improve the customer stickiness, monetizing the user base
- **Model Acceleration & Deployment (100k+ lines of C++)**
 - Experience using various popular Deep Learning Edge Devices for practical applications, such as NVIDIA Xavier, TX2, Nano, HUAWEI Atlas200DK and TDA3x of Texas Instruments (TI)
 - Deployed 80+ models with expertise in the NVIDIA **TensorRT** platform for high-performance inference
 - Familiar with both ARM and x86 architectures, both CPU and GPU instructions for ML Deployment
- **Quantization**
 - Dived into QNNPACK, a Caffe2 8-bit quantization framework, and applied to algorithms within one month of its release from Facebook;
 - Widely used due to its highly efficient performance (**1/4 size, 5x speed, only 1% AP drop**)
 - Given the theoretical analysis of quantization in different tasks (Cls., Seg., Det., Depth) by using the newest INT8 Quantization feature in PyTorch1.3; Deployed 10+ INT8 models based on TorchScript
- **Deep Learning Framework Development**
 - Got familiar with Google XNNPACK, a highly optimized library of neural network inference operators
 - Developed a lite deep learning framework (inference only) for edge devices deployment which needs strong skill on Google FlatBuffers, memory management, topology optimization, etc.
- **Neural Architecture Search (NAS)**
 - Reproduced DARTS, Proxyless NAS (Song Han), Auto-DeepLab (Feifei Li) and Random NAS
 - Extended NAS to **dense image prediction**
- **Training Efficiency**
 - Reduced the training time from 28 GPU-days to 4 GPU-days on GTX-1080Ti by using NVIDIA Data Loading Library (**DALI**) without accuracy reduction (Train ImageNet from scratch)
 - Used **Mixed Precision Training** based on *Tensor Cores* and introduced by Volta Generation of GPUs, to enlarge 8x through put and no accuracy reduction
- **Graph Embedding**
 - **Reproduced** papers related to graph embedding, such as DeepWalk, Node2Vec, and Plantoid

- Proposed a novel semi-supervised approach, **Customized Graph Embedding**, which significantly improved the performance of clustering and representation
- Completed a **first-author paper** in collaboration with MSRA, CMU, UIUC, PKU and SJTU
- **Face Recognition (CNN)**
 - Independently developed a face recognition system for a commercial applications, such as city security
 - **Excellent** performance in both face comparison(95.53% on YTF) and identity verification(99.95%)
- **Photonic Crystals**
 - Completed **research** on Photonic Crystals; analysed results using **MEEP**(MIT open-source package) **simulation** experiments; consolidated findings in the **report**: *Dynamic control of optical pulse delay time*
 - Honored by Tsung-Dao Lee Chinese Research Program with the title "Distinguished Scholar"
 - Selected from the top 3% of applicants to this program supported by Tsung-Dao Lee
- **Misc.** (All of them are implemented for widely use within our company.)
 - Developed a python package for model conversion between DL frameworks (**10k+ lines of .py**)
 - Self-developed an OpenCV (GPU) package to get ride of redundant dependencies (**2k+ lines of .cu**)
 - Python Binding: Created **100+ bindings** of existing C++ code, using C++ code through python API
 - Developed a C++ library for Caffe parser using Google Protocol Buffers (**2k+ lines of .cpp**)
 - Developed a toolchain for CNN debugging with expertise in python features: hook and decorator

Honors & Awards

Outstanding Graduate of SJTU (Top 3%, performances over four years are considered)	2018
Honorable Mention of Mathematical Contest in Modeling(MCM), America	2017
Second Prize in National Undergraduate Electronics Design Contest(Shanghai) (Top 10%)	2017
Tsung-Dao Lee Scholarship (Top 3%, sponsored by the recipient of Nobel Prize in Physics)	2017
Ji Hanbing Alumnus Scholarship (Only 1 in my major, honoring academic excellence)	2017
Liu Yongling Fellowship(Hong Kong) (Only 1 in my major, honoring academic excellence)	2017
Academic Excellence Scholarship (Second Class) of SJTU (Top 3 in my major)	2017
The Merit Student of SJTU (Only 1 per year, comprehensive evaluation)	2016
National Endeavor Fellowship of Shanghai Jiao Tong University (Top 1%, national level)	2016
Third Prize in Texas Instruments(TI) Cup Electronic Design Contest (SJTU) (Top 10)	2016

Conferences, Short-Term Programs, Voluntary and Social Activities

Computing in the 21st Century Conferences & MSRA Faculty Submmmit <i>Invited Audience. One-on-one talk with Kai-Fu Lee and Harry Shum, respectively.</i>	2018
Building Bridges Education Support Program, with Yale U(Organizer), Hong Kong U, Peking U <i>Team Leader. Certificated by the Aixin Foundation Inc. of the United States.</i>	2017
Tsinghua University(THU) Summer Camp: Nano-OptoElectronics Lab <i>Certificate as Outstanding participant by Department of Electronic Engineering, THU</i>	2017
Career & Leadership Development Program <i>Served as Coach. Certified by China Soong Ching Ling Foundation, Liaison Department.</i>	2015
Shanghai International Marathon <i>Volunteer. Served thousands of athletes and running enthusiasts from all over the world.</i>	2014 - 2016

Skills

- **Programming:**Python, C++11/14, CUDA, Verilog/VHDL, HTML, JS, CSS, Java, Neon
- **Dev. SKills:** Linux OS, Git, Modern CMake, GDB, Shell Script(bash, Zsh), Vim, Emacs
- **DL Frameworks:**PyTorch, Caffe, Caffe2; Familiar with Theano, Keras, Tensorflow, MxNet
- **Interests:** Reading, Traveling, Biking, Swimming, Guitar, Badminton