

EDUCATION

- **KTH Royal Institute of Technology** Stockholm, Sweden
Master of Machine Learning; GPA: Excellent *Aug. 2016 – Jun. 2018*
- **Wuhan University** Wuhan, China
Bachelor of Internet of Things; GPA: 3.4 *Sep. 2012 – Jun. 2016*

EXPERIENCE

- **Huawei** Hangzhou, China
Software Engineer *Aug 2018 - Present*
 - **EAI Inference framework:** EAI Inference framework is a pure-C lightweight AI inference framework designed for embedded devices with emphasis on performance. It contains various machine learning algorithms, scientific computing methods as well as a toolkit for model serialization and deserialization. I designed and implemented the model serialization toolkit, various Boosting algorithms, interpolation methods and scientific computing methods, achieved 5x+ performance enhancement than its scipy counterparts.
 - **Mindspore:** Mindspore is a deep learning framework features its source to source automatic differentiation. I worked for both its frontend and its Ascend backend. The frontend is a compiler implemented in C++, it converts python source code into ANF IR and does lots of transformations on it. I designed and implemented a tool which allows users to express IR transformations(passes) in python. As for Ascend backend, I design and implemented several kernels using TIK(TIK is a wrapper language for CCE-C and CCE-C(like CUDA-C) can be compiled into Ascend's ISA).
 - **IC Simulator:** IC Simulator is an IC design verification tool. Essentially it is a compiler for System Verilog(IEEE 1800) which can be used to describe IC designs and testbenches. In this project, we use LLVM and MLIR as our compiler infrastructure. I worked as a feature owner for several features whose job is to make the overall design proposal and to align APIs and progress from parser to runtime. I also worked as a developer for memory allocator which supports tracing-based GC.
 - **Static Timing Analysis tools:** Static timing analysis(STA) is of essential importance through out the physical design stage of a chip. I developed a graph-based analysis tool using Pregel. Pregel is a distributed graph processing framework whose scalability enables processing of gigantic design.
- **KTH Royal Institute of Technology** Stockholm, Sweden
Research Assistant *Aug 2017 - Jun 2018*
 - **Research Assistant - Robotics and Autonomous Systems:** Our group assembled and coded a small autonomous vehicle equipped with multiple sensors and a robotic arm. The final product can drive autonomously in a pre-designed maze and perform grasping tasks in case of target objects' presence. The main components of the software includes sensor fusion, mapping, localization, and planning, and I was mainly responsible for the trajectory planning and the robotic arm's kinematic control.
 - **Research Assistant - Enhance simulated data with GAN:** Deep vision-based robotic grasping system takes a lot of high quality data to train, which is very costly. In this thesis project, a simulated environment is built to collect simulated grasping sessions and an imaged-conditioned GAN is trained to adapt sessions collected in simulated domain to real world domain. For simulated grasping sessions can be easily segmented, the author used this information to eliminate artifacts introduced by the simulated environment and achieved visually realistic results.

PROGRAMMING SKILLS

- **Languages:** C++, C, Python, Matlab, R
- **Technologies:** LLVM, MLIR, Mindspore, TVM, Pybind11, Pregel, ROS
- **Tools:** Git, VS Code, Markdown, Tex, Vim, Hexo

LANGUAGE SKILLS

- **Chinese:** Native
- **English:** Proficient