

WORK EXPERIENCE

- ML Team tech lead, Senior Algorithm Engineer, Magna International, Palo Alto, CA** 2/2020 – present
- **Lead ML team** of 6 software/algorithm engineers in Palo Alto and Canada in development and deployment of the next generation autonomous system on volume production vehicles. Lead cross team collaboration with 5 teams and ~30 members in Michigan, Canada, Germany and India
 - **Lead the in-house Deep Learning Framework** that has following features
 - ❖ PyTorch based modeling/training + Spark based dataset curation
 - ❖ Multi-task learning support: modularized design, supports easy model assembly and hyper-parameter tune-up with pool of available backbone and feature heads, and easy plugin-style module add-on
 - ❖ Multi-data format support: jpeg+json, lmdb, proto and parquet
 - ❖ Multi-language dev-test-deploy environment support with bazel
 - ❖ Containerized and unified dev-test-deploy environment that helps easy deployment and debugging
 - ❖ Production grade tests+lint+doc
 - ❖ Distributed model training/evaluation pipeline with Kubernetes support
 - ❖ Model serving pipeline with gRPC and supports TensorRT/PyTorch/ONNXRuntime backends
 - **Lead Two Multi-tasking model production lines** that deploys on embedded systems on volume production vehicles:
 - ❖ **MagNet:** main production line, target chip: TI TDA4, surrounding view 4 camera solution, 5 feature heads share single backbone:
 - Dynamic object detection
 - Traffic light detection + color/shape classification
 - Trailer detection (bbox+class+angle)
 - Occlusion detection (type+percentage)
 - Semantic segmentation
 - ❖ **TinyMagNet:** Tiny device (confidential), rear camera standalone solution, 2 feature heads share single light weighed backbone
 - Pedestrian detection
 - Occlusion detection (type+percentage)
 - **Lead the autonomous system framework** based on ROS2 as PubSub middleware. Supports multi-container packaging, increased the development speed and system reliability of autonomous vehicle software stack. Organize design and implementation with cross teams collaboration (perception, localization and planning)

- AI Algorithm Engineer, Lyft Level5 - Magna International, Palo Alto, CA** 10/2018 – 1/2020
- One of the main developers and maintainer of the in-house **Machine Learning Framework** with **Tensorflow**. The framework supports 4 teams in development of their ML products.
 - Developed the pipeline **dockerize** the Machine learning models and deploy training on AWS EC2 p3.8xlarge (4 V100) and p3.16xlarge (8 V100) with **Kubernetes**.
 - **Neural Architecture Search (NAS)** via training 300+ of Object Detection models with **Faster RCNN, Single-Shot Detector (SSD)** and **RetinaNet** with various backbones (**Resnet 18/34/50/101, ResNeXt, MobileNet** and their variations), learning schedules, matchers, etc. Found the optimized the structures and improved the mAP for ~10%, Pedestrian and cyclist Recall for ~20% and Precision for ~15%.
 - Created and maintained the **dataset curation pipeline** with **Google BigQuery** and **AWS S3**. Curated and maintained milestone datasets for object detection model training and deployment onto the self-driving fleet.
 - Refactored evaluation pipeline from pytorch to C++ and reduced the evaluation time from ~8hr to 40min
 - Created the **Panorama Object Detection** pipeline by stitching 6 cameras inputs into **360 degree** view image and trained-deployed the **RetinaNet** (variation). Achieved **3X** speed up for total detection latency and dramatically **reduced the duplicated detection and false negatives** on large objects. (Hackathon project)

- AI Algorithm Engineer, Magna International, Auburn Hills, MI** 10/2017 – 9/2018
- Developed model compression algorithms on **YOLOv3-Tiny** (variation) and achieved **1/10** in model size and **3X** speed up.
 - Quantized to int8 with <3% mAP drop **4X** speed up and deployed onto **Xilinx FPGA (Zynq ZCU102)**
 - Established the deployment pipeline and deployed the **quantized (float16)** object detection model (**Caffe/Tensorflow**) to **Nvidia DrivePX2/AGX Xavier** with **TensorRT** and **8 cameras** (max) for **15 FPS** in realtime.

- Graduate Lecturer, The Pennsylvania State University, State College, PA** 05/2017 – 08/2017
Research Assistant, The Pennsylvania State University, State College, PA 08/2014 – 05/2017
Research Assistant, Miami University, Oxford, OH 08/2011 – 08/2013

EDUCATION

- The Pennsylvania State University, State College, PA** 08/2013 – 08/2017
- **Ph.D. Major: Petroleum & Natural Gas Engineering**
 - **Ph.D. Minor: Computational Science**

Miami University, Oxford, OH 08/2011 – 08/2013

 - **MS in Chemical Engineering**

Beijing Jiaotong University, Beijing, China 09/2010 – 07/2012

 - **MS in Optical Engineering**

Beijing Jiaotong University, Beijing, China 09/2006 – 07/2010

 - **BS in Optical Information Science and Technology**

RESEARCH PROJECTS

Dissertation: Automated Neuro-Simulation Protocol for Rate and Pressure Transient Analysis

- Aimed to replace human experts in well time series data analysis by combining **Numerical Computation** with **AI**.
- Combines **Deep Learning (Kernel I)** with **Scientific Computing (Kernel II)** with **C++** to obtain high computational efficiency. Applied modern C++ (14) and coupled with **Python interface**. Built on **NVIDIA Jetson TK1 embedded** development board w/ **CUDA** acceleration.

Kernel I: PSNet: Neural Network Development Framework

- Implemented multiple learning algorithms: **Resilient Back-propagation, Scaled Conjugate Gradient and Levenberg-Marquardt**
- Built Implemented in abstracted matrix operation. Shipped with a unified matrix package interface and can support self-defined/Eigen/Armadillo matrix library. Dramatically reduced the coding/debugging/ computational load.

Kernel II: GNRS: Generalized Numerical Reservoir Simulator

- Designed and implemented a comprehensive and sophisticated numerical simulator for **Computational Fluid Dynamics (CFD)** in porous media
- Support **3D geometry** with complex gridding and coordinating system with multiple phase & component fluid flow
- Applied **automatic differentiation**, dramatically reduced the computational complexity

BOOK CHAPTERS

- Turgay Ertekin, Qian Sun, Jian Zhang, *Reservoir Simulation: Problems and Solutions*, Chapter 1-4, 9. Society of Petroleum Engineers, 2019. ISBN: 978-1-61399-693-5

RECENT INVITED TALKS

- Neuro-Simulation – An Integration of AI & Numerical Optimization in Reservoir Simulation, Smart Field Consortium, School of Earth, Energy & Environmental Sciences, Stanford University, May 2019.

CERTIFICATES

- **Machine Learning (Stanford)** – Coursera
- **Self-Driving Car (Nanodegree)** – Udacity