XIAN SUN

Email: xian.sun@duke.edu Tel: +1 (984)-329-4985

EDUCATION

Duke University, Durham, North Carolina *Aug 2019 - May 2021*
Master of Science in Computer Engineering **GPA:3.97 (Max 4.0)**
Relevant Coursework: Computer Vision, Statistical Machine Learning, Deep Learning, Random Signal and Noise, Vector Space Methods, Systems Programming & Engineering, Software Engineering

Jilin University, Changchun, China *Aug 2015 - June 2019*
Bachelor of Science in Electrical Engineering (rank 1/127) **GPA:3.90 (Max 4.0)**
Relevant Coursework: Advanced Mathematics, Probability and Statistics, Digital Signal Processing, Embedded Systems, Signal and Systems, Linear Algebra, Data Structures and Algorithms

RESEARCH EXPERIENCE

Research Assistant, Computer Vision Lab, Duke University *June 2020 - Present*
Advisor: Dr. Carlo Tomasi
- Built semantic segmentation model (Unet) to segment damages on the herbivory dataset. Developed Leaf-Image-Editor to clean the dataset and designed a copy/paste data augmentation to further improve the accuracy.
- Selected active learning heuristics for the project: Semi-Supervised Active Learning of Optical Flow.

Research Intern, Deep Learning (DL) & ReRAM Group, Duke University *Oct 2020 - Apr 2021*
Advisors: Dr. Krishnendu (Krish) Chakrabarty, Dr. Biresh Joardar
- Saved more than 80 percent of crossbars (128x128) for VGG and ResNet, via extremely sparse neural networks generated by iterative magnitude pruning and index reordering.
- Designed one-shot pruning with thresholds, enabling sparse networks (sparsity > 98%) for AlexNet and VGG.

Research Assistant, Almost Matching Exactly Lab, Duke University *July 2020 - Apr 2021*
Advisors: Dr. Cynthia Rudin, Dr. Sudeepa Roy, Dr. Alexander Volfovsky
- Developed the Python package for Fast Large-scale Almost Matching Exactly on Database (FLAME-DB) Algorithm and R package for Adaptive Hyper Box Matching Algorithm. Both packages support missing data handling, average treatment effect, average treatment effect on treated.
- Improved the speed of the FLAME-DB, with the combination of fixed and adaptive weight matching. The time complexity is reduced to $O(N)$ from $O(N^2)$.

COURSE PROJECTS

Improved Regularization of Convolutional Neural Networks, Duke University *Nov 2020 - Dec 2020*
- Implemented Cutout and Mix-up to address over-fitting problems for ResNet18, 34 and 50.
- Found the best combination of Cutout and Mix-up by grid search on ResNet18-CIFAR10 and transferred it to other datasets like SVHN and Fashion MNIST.

Exploring Probabilistic Classifiers on Binary Classification, Duke University *Mar 2020 - Apr 2020*
- Implemented Bayes classifier based on different assumptions regarding covariance structure of datasets, linear, and logistic discriminant to four different 2-dimensional datasets.

Software Development: RISK Game, Duke University *Feb 2021 - Apr 2021*
- Designed the game server, chat server with non-blocking IO, and state machine in Java.
- Handled concurrency requests with multi-thread programming from multiple users and server-client communication with TCP socket programming.

MANUSCRIPTS

- Shuai Yuan, **Xian Sun**, Hannah Kim, Shuzhi Yu, Carlo Tomasi. Semi-Supervised Active Learning of Optical Flow. Submitted to CVPR 2022
- Neha R. Gupta, Vittorio Orlandi, Chia-Rui Chang, Tianyu Wang, Marco Morucci, Pritam Dey, Thomas J. Howell, **Xian Sun**, Angikar Ghosal, Sudeepa Roy, Cynthia Rudin, Alexander Volfovsky. dame-flame: A Python Library Providing Fast Interpretable Matching for Causal Inference. 2021. arXiv preprint arXiv:2101.01867.

INDUSTRIAL EXPERIENCE

Data Analysis Intern, Neocova, the United States *June 2020- Aug 2020*

- Built a linear regression model that calculates the percentage change in valuation of the model and a binomial logistic model that calculates the probability of positive/negative growth.
- Selected 5 most significant statistically variables with LASSO and 5 demographic variables with Decision Tree from more than 500 variables.

Electrical & Software Intern, Firmenich Aromatics (China) Co. Ltd *April 2019 - June 2019*

- Designed and developed a wireless scoring device that allowed blind employees in food assessment lab to rate products and a software tool that collected and processed data sent by scoring devices simultaneously.
- Improved the lab working efficiency by streamlining the data collecting process.

RELEVANT SKILLS

- **Languages**: Python, Java, C++, R, C, SQL, Markdown, LaTeX, basic HTML/CSS
- **Frameworks**: PyTorch, Scikit-learn, NumPy, Pandas, TensorFlow
- **Software**: Conda, Jupyter Notebook, Tensorboard, PyCharm, IntelliJ, Overleaf, RStudio, PostgreSQL

TEACHING EXPERIENCE

- Instructor in Inspirit AI for AI (high school): AI Scholars and Deep Learning System Design *2021*
- TA to Dr. Carlo Tomasi for CS527 (grad): Computer Vision *2021*
- TA to Dr. Ivan Mura for ECE650K (grad): System Programming & Engineering *2021*
- TA to Dr. Loren Nolte for ECE581 (grad): Random Signal and Noise *2020*
- TA to Dr. Elchanan Solomon for Math216 (undergrad): Linear Algebra & Differential Equations *2020*

HONORS AND AWARDS

National

- China National Scholarship, China Ministry of Education *2017&2018*

University

- ECE Merit Scholarship, Duke University *2020*
- Valedictorian for 2019 Commencement, Jilin University *2019*
- Top Ten Outstanding Student, Jinlin University (10 recipients, academics, service and leadership) *2019*
- Outstanding Volunteer, Jilin University (>600 volunteer hours) *2018*
- CASC Scholarship, China Aerospace Science and Technology Corporation (10 recipients, academics) *2018*