Fan-Keng Sun

fankeng@mit.edu | daikon-sun.github.io | fan-keng-sun | Daikon-Sun

Research Interests

Machine learning and deep learning for sequence modeling, including time series analysis and natural language processing.

Education

Massachusetts Institute of Technology (MIT) *Cambridge, MA*
Ph.D. in Electrical Engineering and Computer Science *09/2019 - PRESENT*

- Courses: Machine Learning, Algorithms for Inference, Advanced Natural Language Processing

National Taiwan University (NTU) *Taipei, Taiwan*
B.S. in Electrical Engineering (major) and Computer Science (minor) *09/2014 - 06/2019*

- GPA: 3.96/4.00, Major GPA: 3.99/4.00

Research Experiences

***Research Assistant*, Statistical Metrology Group**, Prof. Duane Boning *09/2019 - PRESENT*
Machine Learning for Time Series in Manufacturing

***Research Assistant*, Speech Processing and Machine Learning Lab**, Prof. Hung-yi Lee *09/2016 - 08/2019*
*Lifelong Language Learning [**ICLR-20**]*

- Proposed a method that trains a GPT-2 model with both question-answering and language model objectives to tackle the lifelong language learning problem.
- Outperformed previous methods by a large margin yet being simple and extensible.

*Multivariate Time Series (MTS) Forecasting [**Journal track of the ECML/PKDD-19**]*

- Proposed the temporal pattern attention for MTS forecasting, which use CNNs to extract temporal patterns across multiple time steps instead of a single time step as in traditional attention mechanisms.
- Achieved state-of-the-art performance on a wide range of MTS datasets, including polyphonic music notes.

Reviewed paper for ICASSP 2019 and TASLP 2019

***Research Assistant*, Electronic Design Automation Lab**, Prof. Yao-Wen Chang *02/2016 - 12/2018*
*Bivariate Gradient-based Wirelength Model [**DAC-19**]*

- Proposed a novel bivariate gradient-based wirelength model for global placement that combines the advantages of bivariate and multivariate functions.
- Outperformed previous bivariate and state-of-the-art multivariate wirelength models.

*Topology-Matching Bus Routing [**DAC-19**]*

- Proposed and implemented the DAG-based topology-matching bus routing engine and won the top 10 at 2018 ICCAD CAD contest.
- Outperformed all participants of 2018 ICCAD CAD contest, where the 1st place router resulted in 145% higher cost.

*Initial Detailed Routing [**ICCAD-18**]*

- Proposed the multithreaded initial detailed routing engine considering global guides and won the 3rd place at the 2018 ISPD contest.
- Outperformed the winner of 2018 ISPD contest by 23%.

***Undergraduate Researcher*, Speech Processing Lab**, Prof. Lin-shan Lee *09/2017 - 09/2018*
Reinforcing Reinforcement Learning by Rule-based Teacher

- Applied computer vision technique to guide a rule-based Slither.io agent.
- Researched the combination of a rule-based teacher to guide a Slither.io agent by Asynchronous Advantage Actor Critic (A3C) which surpassed rule-based model.

***Software Engineering Intern*, Synopsys, Inc.** *07/2016 - 08/2016*
Single-Layer Global Routing

- Researched and implemented a single-layer global routing algorithm based on a mixture of previous literature and own design.

Publications

(* indicates equal contribution)

1. **Fan-Keng Sun***, Cheng-Hao Ho*, Hung-yi Lee, "LAMOL: LAnguage MOdeling for Lifelong Language Learning", in *International Conference on Learning Representations (ICLR-20)*
2. Shun-Yao Shih*, **Fan-Keng Sun***, Hung-yi Lee, "Temporal Pattern Attention for Multivariate Time Series Forecasting", Journal track of the *European Conference on Machine Learning & Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD-19)*
3. **Fan-Keng Sun**, Yao-Wen Chang, "BiG: A Bivariate Gradient-Based Wirelength Model for Analytical Circuit Placement", in *Proc. of ACM/IEEE Design Automation Conference (DAC-19)*
4. Chen-Hao Hsu, Shao-Chun Hung, Hao Chen, **Fan-Keng Sun**, Yao-Wen Chang, "A DAG-Based Algorithm for Obstacle-Aware Topology-Matching On-Track Bus Routing", in *Proc. of ACM/IEEE Design Automation Conference (DAC-19)*
5. **Fan-Keng Sun**, Hao Chen, Ching-Yu Chen, Chen-Hao Hsu, Yao-Wen Chang, "A Multithreaded Initial Detailed Routing Algorithm Considering Global Routing Guides", in *Proc. of IEEE/ACM International Conference on Computer-Aided Design (ICCAD-18)*

Teachings

***Teaching Assistant*, Algorithm Design and Analysis (Fall 2018)**, Prof. Yun-Nung Chen & Hsu-Chun Hsiao *09/2018 - 01/2019*
***Teaching Assistant*, Machine Learning and Having It Deep and Structured (Spring 2018)**, Prof. Hung-yi Lee *02/2018 - 06/2018*
***Teaching Assistant*, Machine Learning (Fall 2017)**, Prof. Hung-yi Lee *09/2017 - 01/2018*

Honors & Awards

2018 **Outstanding Performance Scholarship**, National Taiwan University
2018 **3rd Place**, Problem A at ICCAD CAD contest
2018 **3rd Place (first pure-undergraduate team in top 3 in 14 years)**, ISPD Contest
2017 **Research Project Grant**, Taiwan Ministry of Science and Technology
2017 **National Technology and Science Scholarship**, CTCI Foundation
2017 **3rd Place**, National Collegiate Programming Contest
2017 **Silver Medal**, ACM ICPC Regional Contest
2016 **1st Place**, ACM ICPC Regional Contest
2016 **1st Place out of 2000 participants from 45 countries**, Calculus World Cup
2016 **2nd Place**, Newcomers for ACM-ICPC Taiwan Online Programming Contest
2012 **Silver Medal**, International Geography Olympiad

Selected Projects

AI Traffic Control System [CTCI Scholarship] *11/2017*

- Designed and implemented a low-cost and real-time traffic signal system on NVIDIA Jetson TK1 using Fast-RCNN to detect the traffic flow and reinforcement learning to train the traffic signal switching interval model.
- Our system is effective on simple traffic simulation, and thus won the 2017 National Technology and Research Scholarship presented by CTCI Foundation.

Solving Multi-Armed Bandits by Upper Confidence Bound (UCB) Algorithms *06/2018*

- Survey several important UCB algorithms: starting from the original UCB, to improved versions (UCBV, improved-UCB), and end at the state-of-the-art method (EUCBV).
- Introduced the lower bound for consistent algorithms and showed the optimality of KL-UCB in special cases.

What does Deep CNN learn? Visualization of Popular Deep CNN Models *03/2017*

- Discussed and compared different methods of visualization for various well-known models in order to gain further insights into the structure and success of CNN.
- Visualization methods includes Activity, Deconvolutional Network, Saliency Map, Deep Generator Network (DGN), and Plug-and-Play Generative Networks.

Extracurricular Activities

***Director*, Academic Department of NTUEE Student Association** *09/2016 - 06/2017*

- Led a team of over 30 students to provide academic services to the 700+ undergraduates, including, but not limited to, the followings:
- EExplore: an event where professors introduce all research areas in EE department to freshmen.
- Lab Intro: a week of continuous lab introduction by the corresponding professor to recruit interested undergraduate researchers.
- Others: online course selection, online textbook bookstore, makerspace, etc.

***Chair*, MakeNTU Makeathon**, [website], [FB fan page], [Recap video] *08/2016 - 02/2017*

- Organized the largest nationwide student Makeathon in Taiwan with 200 participants, 70k USD arrangement, and 60 volunteers.
- Collaborated with the Taipei City and 22 international companies, including Google, Microsoft, Dell, TSMC, Intel, ARM, Asus, etc.

Skills

Natural Languages	Chinese (Mandarin), English, Taiwanese
Programming Languages	Python, C/C++, Shell, Javascript, Matlab, R, LaTeX, Git
Deep Learning Libraries	Tensorflow, PyTorch, Keras