

# Han Shi

PRINCIPAL RESEARCHER

☎ (+852) 59383895 | ✉ shi.han@huawei.com | 🏠 han-shi.github.io | 🎓 Google Scholar

## Education

### Hong Kong University of Science and Technology

PHD IN COMPUTER SCIENCE AND ENGINEERING

Hong Kong

Sep. 2018 - Aug. 2022

### Tsinghua University

BACHELOR IN ELECTRONIC ENGINEERING

Beijing

Sep. 2014 - Jun. 2018

### Tsinghua University

BACHELOR IN ECONOMICS (DUAL)

Beijing

Sep. 2015 - Jun. 2018

## Work Experience

### Huawei Noah Ark's Lab

PRINCIPAL RESEARCHER

Hong Kong

Sep. 2022 - Present

- **Transformer-next (Research Lead)**: building efficient and powerful neural architectures for next generation  
Related Works: *SegLLM, DAPE, DAPE V2, DIM, QuickLLaMA, SJD, SparseBert, BONAS*
- **AI4MATH**: solving formal or informal math problems using Large Language Models  
Related Works: *MetaMath, LEGO-Prover, DT-Solver, FOBAR*

### Media Communication Lab, University of Southern California

RESEARCH ASSISTANT

Los Angeles, California, US

Jun. 2017 - Aug. 2017

## Publications

### DAPE: Data-Adaptive Positional Encoding for Length Extrapolation

C. ZHENG, Y. GAO, [HAN SHI](#), M. HUANG, J. LI, J. XIONG, X. REN, M. NG, X. JIANG, Z. LI, Y. LI

NeurIPS 2024

### Diffusion of thoughts: Chain-of-thought reasoning in diffusion language models

J. YE, S. GONG, L. CHEN, L. ZHENG, J. GAO, [HAN SHI](#), C. WU, X. JIANG, Z. LI, W. BI, L. KONG

NeurIPS 2024

### Forward-Backward Reasoning in Large Language Models for Mathematical Verification

W. JIANG, [HAN SHI](#), L. YU, Z. LIU, Y. ZHANG, Z. LI, J. KWOK

Findings of ACL 2024

### MetaMath: Bootstrap Your Own Mathematical Questions for Large Language Models

L. YU, W. JIANG, [HAN SHI](#)<sup>†</sup>, J. YU, Z. LIU, Y. ZHANG, J. KWOK, Z. LI, A. WELLER, W. LIU

ICLR 2024

Spotlight

### LEGO-Prover: Neural Theorem Proving with Growing Libraries

H. WANG, H. XIN, C. ZHENG, L. LI, Z. LIU, Q. CAO, Y. HUANG, J. XIONG, [HAN SHI](#), E. XIE, J. YIN, Z. LI, H. LIAO, X. LIANG

ICLR 2024

Oral

### DT-Solver: Automated Theorem Proving with Dynamic-Tree Sampling Guided by Proof-level Value Function

H. WANG, Y. YUAN, Z. LIU, J. SHEN, Y. YIN, J. XIONG, E. XIE, [HAN SHI](#), Y. LI, L. LI, J. YIN, Z. LI, X. LIANG

ACL 2023

### DiffFit: Unlocking Transferability of Large Diffusion Models via Simple Parameter-efficient Fine-Tuning

E. XIE, L. YAO, [HAN SHI](#), Z. LIU, D. ZHOU, Z. LIU, J. LI, Z. LI

ICCV 2023

Oral

## GrowCLIP: Data-Aware Automatic Model Growing for Large-scale Contrastive Language-Image Pre-Training

X. DENG, [HAN SHI](#), R. HUANG, C. LI, H. XU, J. HAN, J. KWOK, S. ZHAO, W. ZHANG, X. LIANG

*ICCV 2023*

## Continual Object Detection via Prototypical Task Correlation Guided Gating Mechanism

B. YANG, X. DENG, [HAN SHI](#), C. LI, G. ZHANG, H. XU, S. ZHAO, L. LIN, X. LIANG

*CVPR 2022*

## Revisiting Over-smoothing in BERT from the Perspective of Graph

[HAN SHI](#), J. GAO, H. XU, X. LIANG, Z. LI, L. KONG, S. LEE, J. KWOK

*ICLR 2022*

*Spotlight*

## AutoBERT-Zero: Evolving BERT Backbone from Scratch

J. GAO, H. XU, [HAN SHI](#), X. REN, P.L.H. YU, X. LIANG, X. JIANG, Z. LI

*AAAI 2022*

## SparseBERT: Rethinking the Importance Analysis in Self-attention

[HAN SHI](#), J. GAO, X. REN, H. XU, X. LIANG, Z. LI, J.T. KWOK

*ICML 2021*

## Bridging the Gap between Sample-based and One-shot Neural Architecture Search with BONAS

[HAN SHI](#), R. PI, H. XU, Z. LI, J.T. KWOK, T. ZHANG

*NeurIPS 2020*

## Effective Decoding in Graph Auto-Encoder using Triadic Closure

[HAN SHI](#), H. FAN, J. KWOK

*AAAI 2020*

## Awards

---

2021-22 **RedBird Academic Excellence Award for Continuing PhD Students**, HKUST

*Hong Kong*

2018-22 **Huawei PhD Fellowship**, Huawei

*Hong Kong*

2018-22 **Postgraduate Studentship**, HKUST

*Hong Kong*