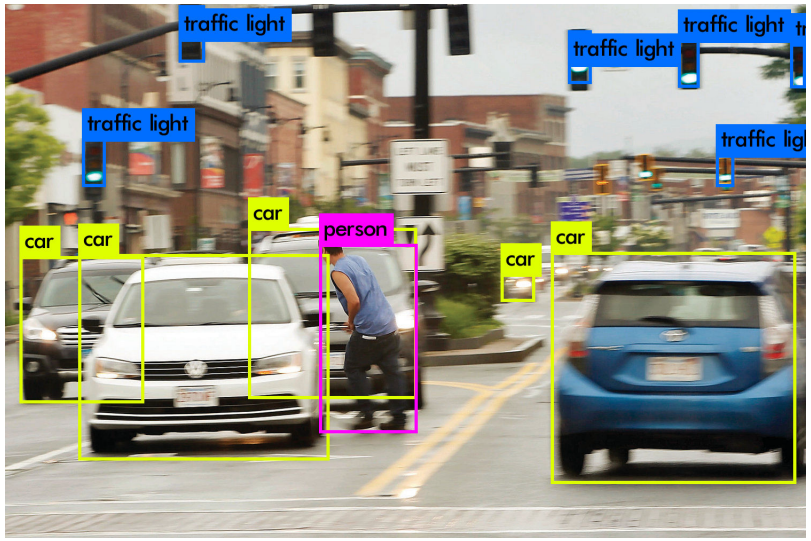
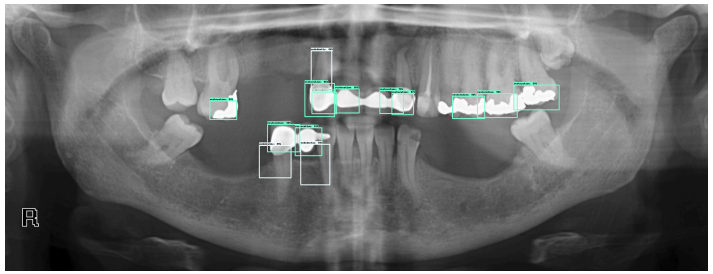
- Shuai Wang (shuaiw@cse.ust.hk)
- Join CSE/HKUST in 2019
- My group: 5 Ph.D. students + 3 MPhil students
- My research interest: Computer Security & Software Engineering

Ensure the reliability of AI models with software engineering methods

The use of AI to support software and system security

Ensure the reliability of AI models with software engineering methods





AI models are employed in security-sensitive scenarios, but they can make mistake…

- Metamorphic Testing for Object Detection Systems – ASE 2020
- Metamorphic Testing for The fairness of Sentiment Analysis Models – IJCAI 2020
- Metamorphic Testing for VQA Models – CVPR 2021*
- Metamorphic Testing for Text-to-SQL Models – ACL 2021*
- Fuzz Testing for Federated Learning Systems – CCS 2021*
- Binary code analysis to exploit/protect compiled DNN models

*Paper currently under review

Metamorphic Testing for Object Detection Systems – ASE 2020

The object detection results over should not change, by inserting an extract object.

- Suppose **no overlapping** with existing objects.

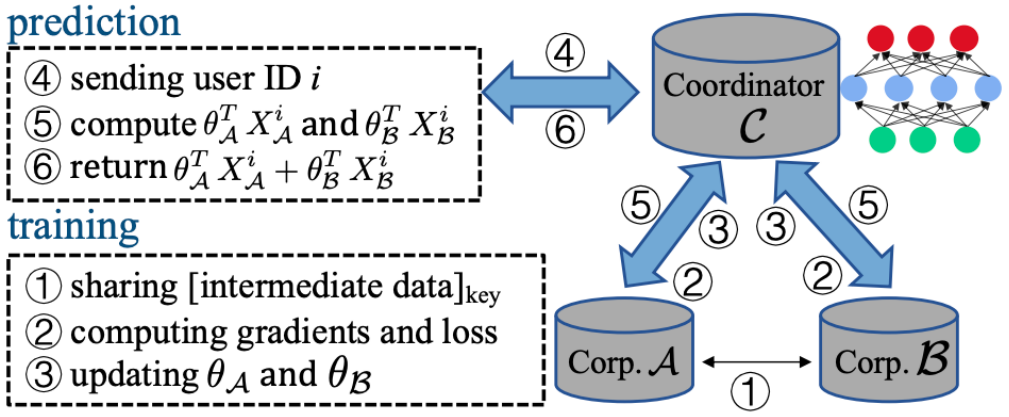
Still can detect two original objects?

If not, then that's a bug!

???

We find tens of thousands of object detection errors in popular (commercial) object detection services.

Fuzz Testing for Federated Learning Systems – CCS 2021



Preserving privacy with VFL

Images Before/After Mutation		Questions	Answers Before/After Mutation	
		How many fruit is here?	24	yes
		Are these real or plastic fruit?	real	yes
		What color are these fruit?	orange	yes
		What is the cat sitting on?	car	cat
		What type of vehicle is there?	car	cat
		Where is the cat?	car	cat
		What route is this bus on?	berkeley bart	1
		How many vans are there?	0	1
		How many trash cans are there?	2	1

But the joint prediction can be dominated by adversarial clients!

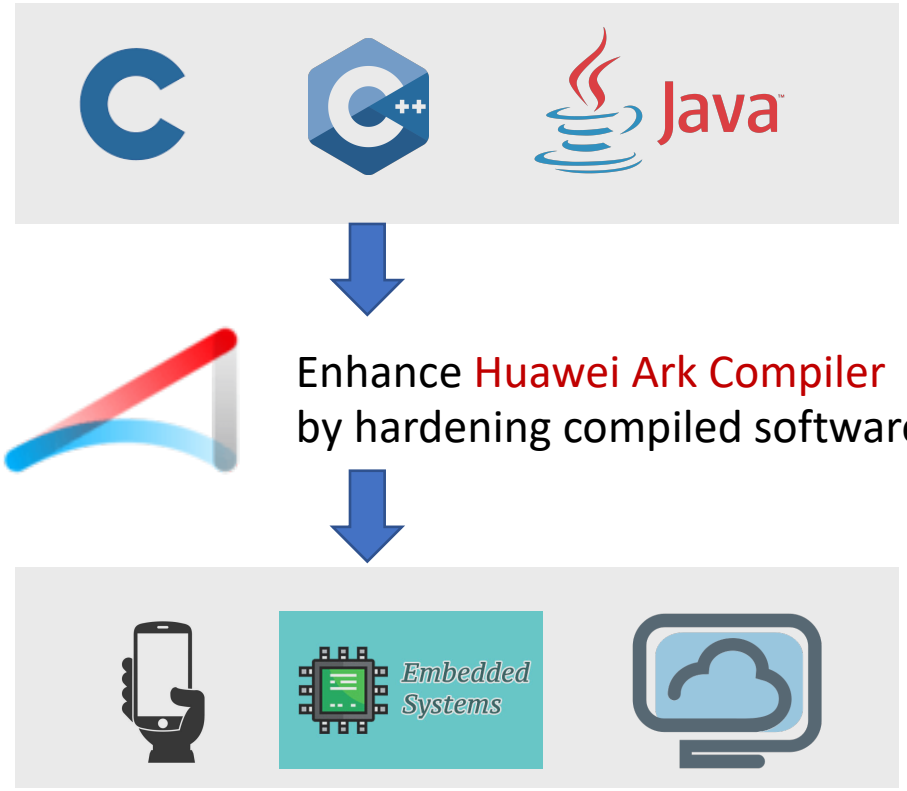
The use of AI to support software and system security

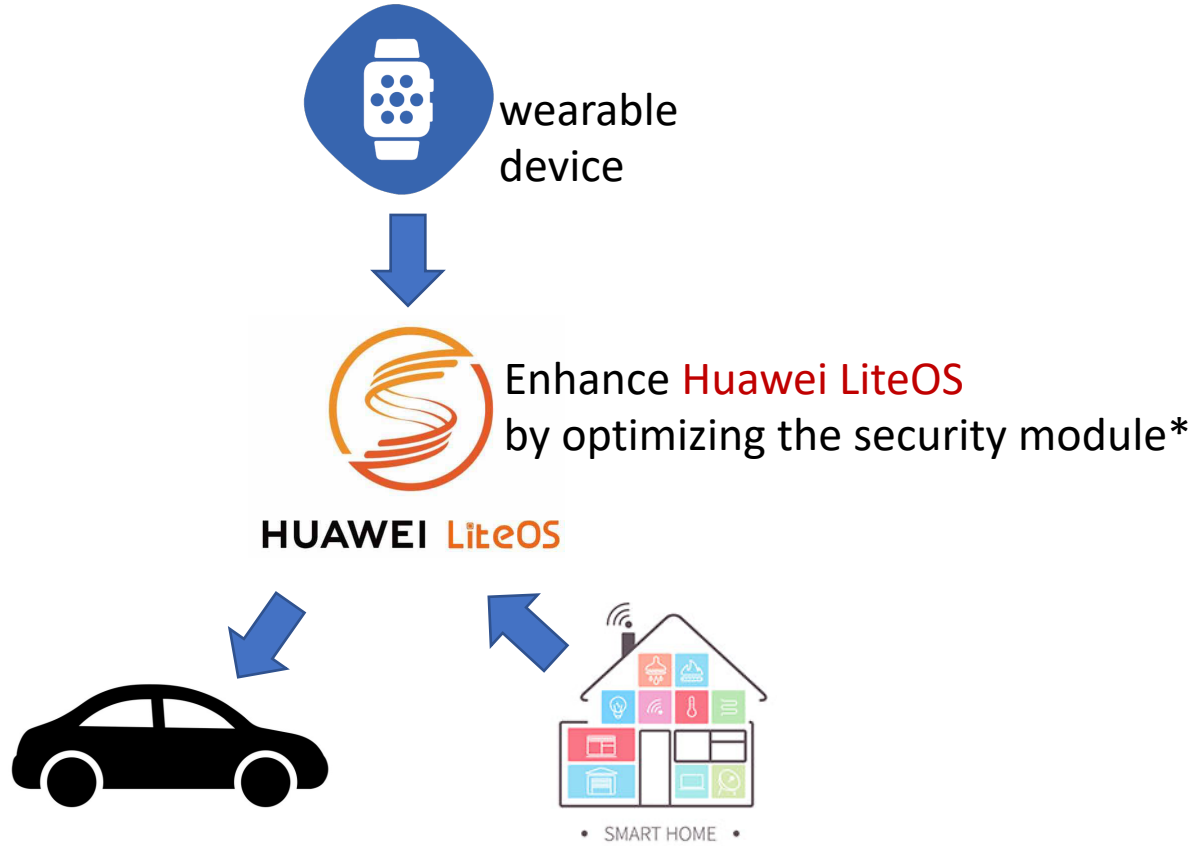
Modern software systems are bloating, but conventional testing/analysis methods are not so scalable or flexible…

- Reinforcement Learning for Software Protection – TDSC 2020
- Generative Models to Exploit Software Systems – ICLR 2021
- Generative Models to Exploit/Protect Software Systems – USENIX Security 2021*
- Adversarial Examples to Exploit Malware Analysis Tools – TIFS 2021*
- Representation Learning + Program Analysis for Software Security Analysis – PLDI 2021*
- Correlation Analysis/Dimension Reduction for More Efficient Software Protection – OSDI 2021*
- Reinforcement Learning for More Efficient Software Protection
- Reinforcement Learning for Software Fuzz Testing

*Paper currently under review

Real-World Impact





Technically, ALL software in the Huawei ecosystem will be secured by research products of my group.

Research Grants Council
of Hong Kong
香港 研究資助局

HUAWEI

Tencent 腾讯

*Technical details cannot be disclosed due to NDA