杨定澄

86-18810229015(mobile) ⋄ ydc19@mails.tsinghua.edu.cn

清华大学

教育经历

博士生, 清华大学 *2019.7 -*

计算机科学与技术专业

本科, 清华大学 *2015.9 - 2019.7*

计算机科学与技术专业

研究兴趣

研究领域包括 AI 安全、模型压缩、AI for Science。主要从事 AI 安全相关研究，如对抗样本的生成与防御、AIGC 的生成与检测。最近我对生成式大模型特别是 Diffusion 模型相关的安全问题有较大兴趣。

论文发表

Boosting the adversarial transferability of surrogate models with dark knowledge

- **Dingcheng Yang**, Zihao Xiao, Wenjian Yu.
- International Conference on Tools with Artificial Intelligence (**ICTAI**), 2023.
- Arxiv link: https://arxiv.org/abs/2206.08316.
- Code link: https://github.com/ydc123/Dark_Surrogate_Model.

Generating Adversarial Examples with Better Transferability via Masking Unimportant Parameters of Surrogate Model

- **Dingcheng Yang**, Wenjian Yu, Zihao Xiao, Jiaqi Luo.
- International Joint Conference on Neural Networks (**IJCNN**), 2023.
- Paper link: https://ieeexplore.ieee.org/document/10191679.
- Code link: https://github.com/ydc123/MUP_Attack.

CNN-Cap: Effective Convolutional Neural Network Based Capacitance Models for Interconnect Capacitance Extraction

- **Dingcheng Yang**, Haoyuan Li, Wenjian Yu, Yuanbo Guo, Wenjie Liang.
- ACM Transactions on Design Automation of Electronic Systems (**TODAES**), 2022.
- Paper link: https://dl.acm.org/doi/abs/10.1145/3564931.
- Code link: https://github.com/ydc123/CNNCap.

DP-Nets: Dynamic programming assisted quantization schemes for DNN compression and acceleration

- **Dingcheng Yang**, Wenjian Yu, Xiangyun Ding, Ao Zhou, Xiaoyi Wang.
- Integration, the VLSI Journal (**Integration**), 2022.

- Paper link: https://dl.acm.org/doi/abs/10.1016/j.vlsi.2021.10.002.

CNN-Cap: Effective convolutional neural network based capacitance models for full-chip parasitic extraction

- **Dingcheng Yang**, Wenjian Yu, Yuanbo Guo and Wenjie Liang.
- International Conference on Computer-Aided Design (**ICCAD**), Munich, Germany, 2021.
- Paper link: https://dl.acm.org/doi/abs/10.1109/ICCAD51958.2021.9643461.
- Code link: https://github.com/ydc123/CNNCap.

Dynamic Programming Assisted Quantization Approaches for Compressing Normal and Robust DNN Models

- **Dingcheng Yang**, Wenjian Yu, Haoyuan Mu, Gary Yao.
- Asia and South Pacific Design Automation Conference (**ASPDAC**), 2021.
- Paper link: https://dl.acm.org/doi/abs/10.1145/3394885.3431538.

Optimal Algorithm for Profiling Dynamic Arrays with Finite Values

- **Dingcheng Yang**, Wenjian Yu, Junhui Deng, Shenghua Liu.
- International Conference on Extending Database Technology (**EDBT**), 2019.
- Paper link: http://openproceedings.org/2019/conf/edbt/EDBT19_paper_279.pdf.

Training better CNN models for 3-D capacitance extraction with neural architecture search

- Haoyuan Li*, **Dingcheng Yang*** and Wenjian Yu. (*Equal contribution)
- Design, Automation and Test in Europe Conference (**DATE**), 2024.

Pose2Seg: Detection Free Human Instance Segmentation

- Songhai Zhang, Ruilong Li, Xin Dong, Paul L. Rosin, Zixi Cai, Han Xi, **Dingcheng Yang**, Haozhi Huang, Shimin Hu.
- IEEE conference on computer vision and pattern recognition (**CVPR**), 2019.
- Paper link: https://ieeexplore.ieee.org/abstract/document/8953934.
- Code link: https://github.com/liruilong940607/Pose2Seg.

预印本

RobFR: Benchmarking Adversarial Robustness on Face Recognition

- Xiao Yang, **Dingcheng Yang**, Yinpeng Dong, Hang Su, Wenjian Yu, Jun Zhu.
- Paper link: https://arxiv.org/abs/2007.04118.
- Code link: https://github.com/ShawnXYang/Face-Robustness-Benchmark.

专利

基于多先验的黑盒对抗测试样本生成方法及装置。喻文健，**杨定澄**。

对抗样本生成方法、装置、介质和计算设备。萧子豪，董胤蓬，**杨定澄**。

深度神经网络的模型压缩方法及系统 (pending), 喻文健，**杨定澄**。

对抗样本图像生成方法及装置、电子设备和存储介质 (pending), 喻文健，**杨定澄**。

模型蒸馏方法、装置、电子设备和可读存储介质 (pending), 喻文健，**杨定澄**。

实习经历

腾讯（Tencent） 2018-07-01 - 2018-08-31

- 2018 年 7 月-8 月，在腾讯公司，担任产品开发组暑期实习生，参与腾讯微校场馆预约管理系统的后端开发。
- 在实习期间，产品成功上线。

瑞莱智慧（RealAI） 2019-11-01 - 至今

- 在瑞莱智慧有限公司担任算法实习生，研究和 AI 相关的安全问题。
- 期间为公司产出多项竞赛冠军、论文、专利。
- 曾获公司年度最佳实习生奖。

服务

参与承办AAAI 2022 Workshop Adversarial Machine Learning and Beyond以及对应比赛AAAI-2022 安全 AI 挑战者计划第八期：以数据为中心的鲁棒机器学习。

参与承办AISC 人工智能安全大赛。

审稿工作: ECAI2023, IJCNN2023, ECCV2022, AAAIW2022, ECCVW2022。

所获奖项

全国青少年信息学奥林匹克竞赛 (NOI) 金牌 (公示)。2014.7

GeekPwn CAAD (Competition on Adversarial Attacks and Defenses) CTF competitions 冠军。(B 站视频地址)。对抗样本生成竞赛，我是全场唯一攻破商用人脸识别模型的选手。2019.10

GeekPwn DeepFake competitions 冠军。(B 站视频地址)。比赛内容为人脸的深度伪造与检测。2020.10

EDAthon Contest 冠军。2020.8

RealAI 年度最佳实习生。2022.1

龙湖奖学金风采奖。2022.9

清华之友-浦口英才二等奖学金。2022.10

清华之友-浦口英才二等奖学金。2021.10

EDA Elite Challenge 二等奖。2020.11

IJCAI-19 Alibaba Adversarial AI Challenge 第九名。2019.5

AI Challenger: Human Skeletal System Keypoints Detection Competition 第六名。2017.12

技能

编程: Python, C/C++, Matlab, Java, php, Verilog。

软件与工具: PyTorch, Tensorflow, NumPy, Flask, Lumen。

语言: 中文（母语）, 英语（流利）。