

Yinpeng Dong

Department of Computer Science and Technology, Tsinghua University
FIT building 1-509, Beijing, China, 100084
<http://ml.cs.tsinghua.edu.cn/~yinpeng>

Contact	Phone: (+86) 18603303421 Email: dongyinpeng@mail.tsinghua.edu.cn ; dongyinpeng@gmail.com
Work Experience	Department of Computer Science and Technology Tsinghua University, Beijing, China Postdoctoral Researcher , collaborated with Prof. Jun Zhu 2022.01 -
Education	Department of Computer Science and Technology Tsinghua University, Beijing, China Ph.D , advised by Prof. Jun Zhu 2017.09 - 2022.01
	Department of Computer Science and Technology Tsinghua University, Beijing, China Bachelor of Engineering GPA: 94.4/100; Rank: 2/107 2013.08 - 2017.06
	Robotic Institute Carnegie Mellon University, Pittsburgh, US Visiting Student 2016.06 - 2016.09
Selected Awards	Tsinghua Outstanding Postdoctoral Researcher (Top 10 in Tsinghua) 2023.07 CCF Outstanding Doctoral Dissertation Award (Top 10 in China) 2022.12 National Postdoctoral Innovative Talents Support Program 2022.06 Shuimu Tsinghua Scholar Program 2022.01 Beijing Outstanding Graduates 2022.01 ByteDance Scholars Program (Top 10 in China) 2020.11 Baidu Fellowship (Top 10 Worldwide) 2020.01 Microsoft Research Asia (MSRA) Fellowship (Top 12 in Asia) 2019.11 VALSE Annual Outstanding Student Paper Award (Top 3 in China) 2019.04 CCF-CV Academic Emerging Award (Top 3 in China) 2018.11
Publications	(* indicates equal contribution, † indicates corresponding author) Robust Classification via a Single Diffusion Model Huanran Chen, Yinpeng Dong [†] , Zhengyi Wang, Xiao Yang, Chengqi Duan, Hang Su, Jun Zhu [†] <i>International Conference on Machine Learning (ICML), 2024</i> Efficient Black-box Adversarial Attacks via Bayesian Optimization Guided by a Function Prior Shuyu Cheng, Yibo Miao, Yinpeng Dong [†] , Xiao Yang, Xiao-Shan Gao, Jun Zhu <i>International Conference on Machine Learning (ICML), 2024</i>

Machine Vision Therapy: Multimodal Large Language Models Can Enhance Visual Robustness via Denoising In-Context Learning

Zhuo Huang, Chang Liu, **Yinpeng Dong**, Hang Su, Shibao Zheng, Tongliang Liu
International Conference on Machine Learning (ICML), 2024

Toward Availability Attacks in 3D Point Clouds

Yifan Zhu, Yibo Miao, **Yinpeng Dong**, Xiao-Shan Gao
International Conference on Machine Learning (ICML), 2024

Exploring the Transferability of Visual Prompting for Multimodal Large Language Models (**Highlight**)

Yichi Zhang, **Yinpeng Dong**[†], Siyuan Zhang, Tianzan Min, Hang Su[†], Jun Zhu
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024

Towards Transferable Targeted 3D Adversarial Attack in the Physical World

Yao Huang, **Yinpeng Dong**[†], Shouwei Ruan, Xiao Yang, Hang Su, Xingxing Wei[†]
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024

Focus on Hiders: Exploring Hidden Threats for Enhancing Adversarial Training

Qian Li, Yuxiao Hu, **Yinpeng Dong**, Dongxiao Zhang, Yuntian Chen
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024

Rethinking Model Ensemble in Transfer-based Adversarial Attacks

Huanran Chen, Yichi Zhang, **Yinpeng Dong**[†], Xiao Yang, Hang Su, Jun Zhu[†]
International Conference on Learning Representations (ICLR), 2024

Embodied Active Defense: Leveraging Recurrent Feedback to Counter Adversarial Patches

Lingxuan Wu, Xiao Yang, **Yinpeng Dong**, Liuwei Xie, Hang Su, Jun Zhu
International Conference on Learning Representations (ICLR), 2024

Learning Sample Difficulty from Pre-trained Models for Reliable Prediction

Peng Cui, Dan Zhang, Zhijie Deng, **Yinpeng Dong**, Jun Zhu
Advances in Neural Information Processing Systems (NeurIPS), 2023

Towards Viewpoint-Invariant Visual Recognition via Adversarial Training

Shouwei Ruan, **Yinpeng Dong**, Hang Su, Jianteng Peng, Ning Chen, Xingxing Wei
International Conference on Computer Vision (ICCV), 2023

Root Pose Decomposition Towards Generic Non-rigid 3D Reconstruction with Monocular Videos

Yikai Wang, **Yinpeng Dong**, Fuchun Sun, Xiao Yang
International Conference on Computer Vision (ICCV), 2023

Text-to-Image Diffusion Models can be Easily Backdoored through Multimodal Data Poisoning (**Oral**)

Shengfang Zhai, **Yinpeng Dong**[†], Qingni Shen, Shi Pu, Yuejian Fang[†], Hang Su
ACM International Conference on Multimedia (MM), 2023

GNOT: A General Neural Operator Transformer for Operator Learning

Zhongkai Hao, Zhengyi Wang, Hang Su, Chengyang Ying, **Yinpeng Dong**, Songming Liu, Ze Cheng, Jian Song, Jun Zhu
International Conference on Machine Learning (ICML), 2023

Benchmarking Robustness of 3D Object Detection to Common Corruptions in Autonomous Driving

Yinpeng Dong, Caixin Kang, Jinlai Zhang, Zijian Zhu, Yikai Wang, Xiao Yang, Hang Su, Xingxing Wei, Jun Zhu

IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023

Towards Effective Adversarial Textured 3D Meshes on Physical Face Recognition (**Highlight**)

Xiao Yang, Chang Liu, Longlong Xu, Yikai Wang, **Yinpeng Dong**[†], Ning Chen, Hang Su, Jun Zhu[†]

IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023

Understanding the Robustness of 3D Object Detectors with Bird's-Eye-View Representations in Autonomous Driving

Zijian Zhu, Yichi Zhang, Hai Chen, **Yinpeng Dong**[†], Shu Zhao, Wenbo Ding, Jiachen Zhong, Shibao Zheng[†]

IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023

Compacting Binary Neural Networks by Sparse Kernel Selection

Yikai Wang, Wenbing Huang, **Yinpeng Dong**, Fuchun Sun, Anbang Yao

IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023

ViewFool: Evaluating the Robustness of Visual Recognition to Adversarial Viewpoints

Yinpeng Dong, Shouwei Ruan, Hang Su, Caixin Kang, Xingxing Wei, Jun Zhu

Advances in Neural Information Processing Systems (NeurIPS), 2022

Pre-trained Adversarial Perturbations

Yuanhao Ban, **Yinpeng Dong**[†]

Advances in Neural Information Processing Systems (NeurIPS), 2022

Isometric 3D Adversarial Examples in the Physical World

Yibo Miao, **Yinpeng Dong**[†], Jun Zhu, Xiao-Shan Gao[†]

Advances in Neural Information Processing Systems (NeurIPS), 2022

Boosting Transferability of Targeted Adversarial Examples via Hierarchical Generative Networks

Xiao Yang, **Yinpeng Dong**, Tianyu Pang, Hang Su, Jun Zhu

European Conference on Computer Vision (ECCV), 2022

AutoDA: Automated Decision-based Iterative Adversarial Attacks

Qi-An Fu, **Yinpeng Dong**, Hang Su, Jun Zhu, Chao Zhang

31st USENIX Security Symposium (USENIX Security '22), 2022

GSmooth: Certified Robustness against Semantic Transformations via Generalized Randomized Smoothing

Zhongkai Hao, Chengyang Ying, **Yinpeng Dong**, Hang Su, Jian Song, Jun Zhu

International Conference on Machine Learning (ICML), 2022

Two Coupled Rejection Metrics Can Tell Adversarial Examples Apart

Tianyu Pang, Huishuai Zhang, Di He, **Yinpeng Dong**, Hang Su, Wei Chen, Jun Zhu, Tie-Yan Liu

IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022

Exploring Memorization in Adversarial Training

Yinpeng Dong, Ke Xu, Xiao Yang, Tianyu Pang, Zhijie Deng, Hang Su, Jun Zhu
International Conference on Learning Representations (ICLR), 2022

Query-Efficient Black-box Adversarial Attacks Guided by a Transfer-based Prior
Yinpeng Dong*, Shuyu Cheng*, Tianyu Pang, Hang Su, Jun Zhu
IEEE Transaction on Pattern Analysis and Machine Intelligence (TPAMI), 2021

Accumulative Poisoning Attacks on Real-time Data
Tianyu Pang, Xiao Yang, **Yinpeng Dong**, Hang Su, Jun Zhu
Advances in Neural Information Processing Systems (NeurIPS), 2021

Black-box Detection of Backdoor Attacks with Limited Information and Data
Yinpeng Dong, Xiao Yang, Zhijie Deng, Tianyu Pang, Zihao Xiao, Hang Su, Jun Zhu
International Conference on Computer Vision (ICCV), 2021

Towards Face Encryption by Generating Adversarial Identity Masks
Xiao Yang, **Yinpeng Dong**, Tianyu Pang, Hang Su, Jun Zhu, Yuefeng Chen, Hui Xue
International Conference on Computer Vision (ICCV), 2021

Improving Transferability of Adversarial Patches on Face Recognition with Generative Models
Zihao Xiao, Xianfeng Gao, Chilin Fu, **Yinpeng Dong**, Wei Gao, Xiaolu Zhang, Jun Zhou, Jun Zhu
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021

Bag of Tricks for Adversarial Training
Tianyu Pang, Xiao Yang, **Yinpeng Dong**, Hang Su, Jun Zhu
International Conference on Learning Representations (ICLR), 2021

Adversarial Distributional Training for Robust Deep Learning
Yinpeng Dong*, Zhijie Deng*, Tianyu Pang, Hang Su, Jun Zhu
Advances in Neural Information Processing Systems (NeurIPS), 2020

Understanding and Exploring the Network with Stochastic Architectures
Zhijie Deng, **Yinpeng Dong**, Shifeng Zhang, Jun Zhu
Advances in Neural Information Processing Systems (NeurIPS), 2020

Boosting Adversarial Training with Hypersphere Embedding
Tianyu Pang*, Xiao Yang*, **Yinpeng Dong**, Kun Xu, Hang Su, Jun Zhu
Advances in Neural Information Processing Systems (NeurIPS), 2020

Benchmarking Adversarial Robustness on Image Classification (**Oral**)
Yinpeng Dong, Qi-An Fu, Xiao Yang, Tianyu Pang, Hang Su, Zihao Xiao, Jun Zhu
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020

Rethinking Softmax Cross-Entropy Loss for Adversarial Robustness
Tianyu Pang, Kun Xu, **Yinpeng Dong**, Chao Du, Ning Chen, Jun Zhu
International Conference on Learning Representations (ICLR), 2020

Improving Black-box Adversarial Attacks with a Transfer-based Prior
Shuyu Cheng*, **Yinpeng Dong***, Tianyu Pang, Hang Su, Jun Zhu
Advances in Neural Information Processing Systems (NeurIPS), 2019

Evading Defenses to Transferable Adversarial Examples by Translation-Invariant Attacks (**Oral**)

Yinpeng Dong, Tianyu Pang, Hang Su, Jun Zhu

IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019

Efficient Decision-based Black-box Adversarial Attacks on Face Recognition

Yinpeng Dong, Hang Su, Baoyuan Wu, Zhifeng Li, Wei Liu, Tong Zhang, Jun Zhu

IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019

Stochastic Quantization for Learning Accurate Low-bit Deep Neural Networks

Yinpeng Dong, Renkun Ni, Jianguo Li, Yurong Chen, Hang Su, Jun Zhu

International Journal of Computer Vision (IJCV), 2019

Composite Binary Decomposition Networks

You Qiaoben, Zheng Wang, Jianguo Li, **Yinpeng Dong**, Yu-Gang Jiang, Jun Zhu

The Thirty-Third AAAI Conference on Artificial Intelligence (AAAI), 2019

Towards Robust Detection of Adversarial Examples (**Spotlight**)

Tianyu Pang, Chao Du, **Yinpeng Dong**, Jun Zhu

Advances in Neural Information Processing Systems (NeurIPS), 2018

Boosting Adversarial Attacks with Momentum (**Spotlight**)

Yinpeng Dong, Fangzhou Liao, Tianyu Pang, Hang Su, Jun Zhu, Xiaolin Hu, Jianguo Li

IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018

Defense against Adversarial Attacks Using High-Level Representation Guided Denoiser

Fangzhou Liao*, Ming Liang*, **Yinpeng Dong**, Tianyu Pang, Jun Zhu, Xiaolin Hu

IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018

Learning Visual Knowledge Memory Networks for Visual Question Answering

Zhou Su, Chen Zhu, **Yinpeng Dong**, Dongqi Cai, Yurong Chen, Jianguo Li

IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018

Learning Accurate Low-Bit Deep Neural Networks with Stochastic Quantization (**Oral, Best Paper Nomination**)

Yinpeng Dong, Renkun Ni, Jianguo Li, Yurong Chen, Jun Zhu, Hang Su

British Machine Vision Conference (BMVC), 2017

Forecast Plausible Paths in Crowd Scenes

Hang Su, Jun Zhu, **Yinpeng Dong**, Bo Zhang

International Joint Conference on Artificial Intelligence (IJCAI), 2017

Improving Interpretability of Deep Neural Networks with Semantic Information

Yinpeng Dong, Hang Su, Jun Zhu, Bo Zhang

IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017

Crowd Scene Understanding with Coherent Recurrent Neural Networks

Hang Su, **Yinpeng Dong**, Jun Zhu, Haibin Ling, Bo Zhang

International Joint Conference on Artificial Intelligence (IJCAI), 2016

Competitions

The 1st place in the Adversarial Robustness of Deep Learning track of 2022 International Algorithm Case Competition

2022.12

The 1st place in GeekPwn DeepFake competition (Shanghai) 2020.10

The 1st places in GeekPwn CAAD CTF and Adversarial Patch competitions (Shanghai) 2019.10

The 2nd place in the Untargeted Attack track of NeurIPS 2018 Adversarial Vision Challenge 2018.11

The 2nd places in Targeted Attack track, Defense track, and the 3rd place in Non-targeted Attack track of GeekPwn CAAD competition 2018.9

The 1st place in GeekPwn CAAD CTF competition (Las Vegas) 2018.8

The 1st places in NeurIPS 2017 Adversarial Attacks and Defenses 2017.10

Services

Organizer for:

ICCV 2023 Workshop on Adversarial Robustness in the Real World
ECCV 2022 Workshop on Adversarial Robustness in the Real World
AAAI 2022 Workshop on Adversarial Machine Learning and Beyond
ICML 2021 Workshop on A Blessing in Disguise: The Prospects and Perils of Adversarial Machine Learning
ICCV 2021 Workshop on Adversarial Robustness in the Real World
CVPR 2021 Workshop on Adversarial Machine Learning in Real-World Computer Vision Systems and Online Challenges (AML-CV)

Reviewer for:

TPAMI 2019-2024; **IJCV** 2021-2023; **TIP** 2019-2021; **NeurIPS** 2016, 2019-2023; **ICML** 2019-2024; **CVPR** 2019-2024; **ICLR** 2020-2024; **ICCV** 2019, 2021, 2023

Teaching

Lecturer in *CCF ADL140: Robust Machine Learning* 2023.06

Head TA in *Statistical Machine Learning*, instructed by Prof. Jun Zhu 2019 Spring