Research Landscape on Robust Perception

Felix Juefei-Xu
xujuefei.com

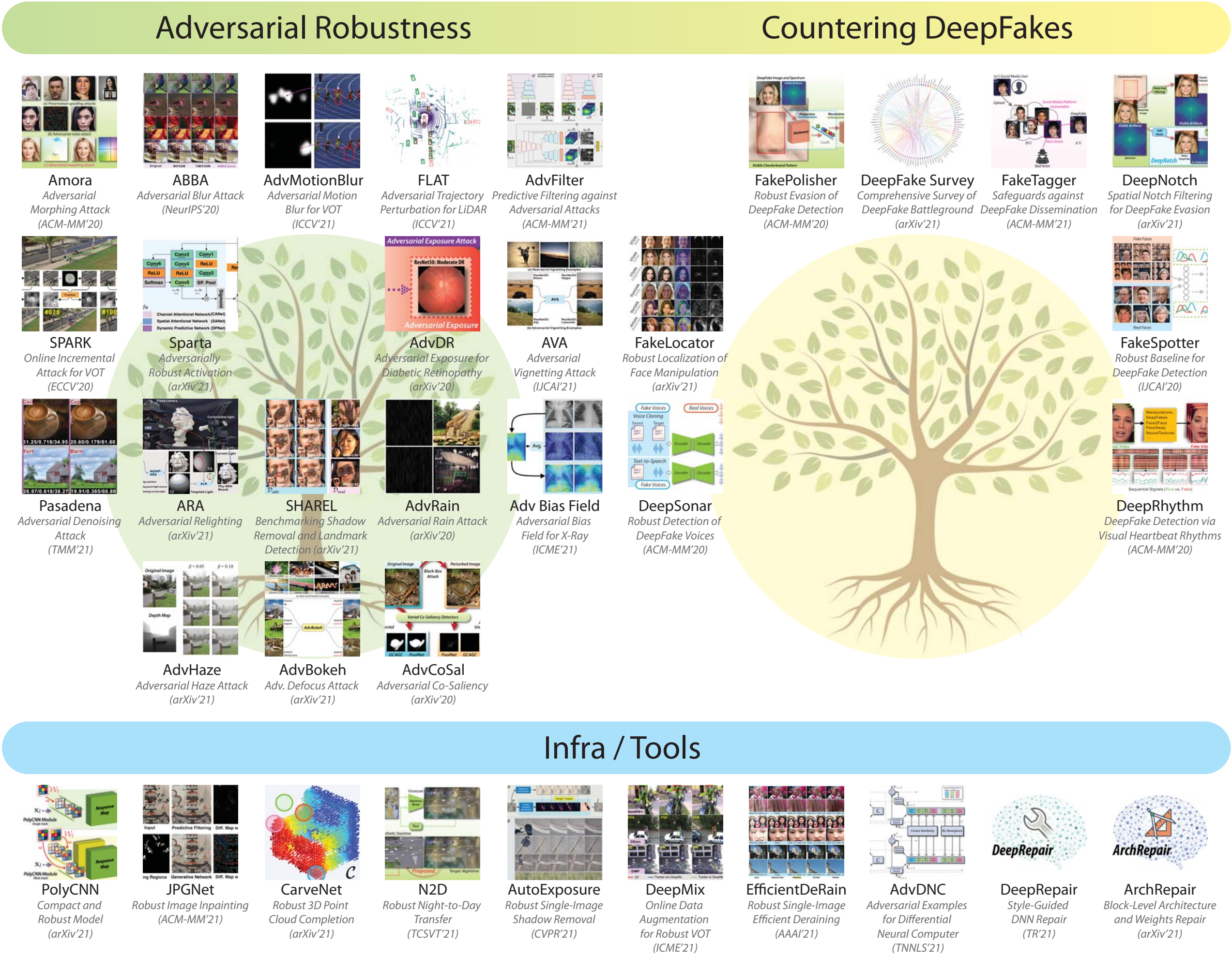
Adversarial Robustness
Countering DeepFakes
Amora
Adversarial Morphing Attack
(ACM-MM'20)
ABBA
Adversarial Blur Attack
(NeurIPS'20)
AdvMotionBlur
Adversarial Motion Blur for VOT
(ICCV'21)
FLAT
Adversarial Trajectory Perturbation for LiDAR
(ICCV'21)
AdvFilter
Predictive Filtering against Adversarial Attacks
(ACM-MM'21)
FakePolisher
Robust Evasion of DeepFake Detection
(ACM-MM'20)
DeepFake Survey
Comprehensive Survey of DeepFake Battleground
(arXiv'21)
FakeTagger
Safeguards against DeepFake Dissemination
(ACM-MM'21)
DeepNotch
Spatial Notch Filtering for DeepFake Evasion
(arXiv'21)
SPARK
Online Incremental Attack for VOT
(ECCV'20)
Sparta
Adversarially Robust Activation
(arXiv'21)
AdvDR
Adversarial Exposure for Diabetic Retinopathy
(arXiv'20)
AVA
Adversarial Vignetting Attack
(IJCAI'21)
FakeLocator
Robust Localization of Face Manipulation
(arXiv'21)
FakeSpotter
Robust Baseline for DeepFake Detection
(IJCAI'20)
Pasadena
Adversarial Denoising Attack
(TMM'21)
ARA
Adversarial Relighting
(arXiv'21)
SHAREL
Benchmarking Shadow Removal and Landmark Detection (arXiv'21)
AdvRain
Adversarial Rain Attack
(arXiv'20)
Adv Bias Field
Adversarial Bias Field for X-Ray
(ICME'21)
DeepSonar
Robust Detection of DeepFake Voices
(ACM-MM'20)
DeepRhythm
DeepFake Detection via Visual Heartbeat Rhythms
(ACM-MM'20)
AdvHaze
Adversarial Haze Attack
(arXiv'21)
AdvBokeh
Adv. Defocus Attack
(arXiv'21)
AdvCoSal
Adversarial Co-Saliency
(arXiv'20)
Infra / Tools
PolyCNN
Compact and Robust Model
(arXiv'21)
JPGNet
Robust Image Inpainting
(ACM-MM'21)
CarveNet
Robust 3D Point Cloud Completion
(arXiv'21)
N2D
Robust Night-to-Day Transfer
(TCSVT'21)
AutoExposure
Robust Single-Image Shadow Removal
(CVPR'21)
DeepMix
Online Data Augmentation for Robust VOT
(ICME'21)
EfficientDeRain
Robust Single-Image Efficient Deraining
(AAAI'21)
AdvDNC
Adversarial Examples for Differential Neural Computer
(TNNLS'21)
DeepRepair
Style-Guided DNN Repair
(TR'21)
ArchRepair
Block-Level Architecture and Weights Repair
(arXiv'21)

Figure 1. A snippet of my latest research landscape on robust perception as of December of 2021.

My latest research landscape on robust perception is shown in Figure 1. My research in general is focused on a fuller understanding of deep learning where I am actively exploring new methods in deep learning that are statistically efficient and adversarially robust [23,24]. Through these efforts, I am investigating under what conditions deep learning starts to work and under what conditions deep learning starts to fail.

I am now heavily involved in pushing the boundary of robust vision and safe AI. As shown in Figure 1, the first research direction is on *adversarial robustness* where I have created high-realism degradation-mimetic adversarial attacks such as adversarial motion blur [12] / defocus blur [21], rain [42] / haze [8], relighting [10] / exposure [7, 34] / vignetting [35] / shadow [5], *etc.*, and then devising tools to defend against them through an attempt to situate both the natural corruption robustness and the adversarial robustness problems under the same dome [3, 4, 9, 11, 17, 18, 27, 36].

The second research direction in on *countering malicious DeepFakes*, where I have created cunning DeepFakes [26] that are detection-evasive [20, 22] to foster the development of next-generation robust detectors of real-world video/audio DeepFakes [19, 32, 37–39].

The third research direction is on *infrastructural tools* for robust perception, where I have sought robust learners by creating novel neural architectures and computational modules that are statistically efficient and adversarially robust [1, 13, 23–25], as well as multifarious task-specific processing tools [2, 6, 14–16] and deep neural network robustness gauging and repairing techniques [28–31, 33, 40, 41].

References

[1] Ramzi Abiantun, Felix Juefei-Xu, Utsav Prabhu, and Marios Savvides. SSR2: Sparse Signal Recovery for Single-Image Super-Resolution on Faces with Extreme Low Resolutions. *Pattern Recognition*, 2019.

[2] Alvin Chan, Lei Ma, Felix Juefei-Xu, Yew Soon Ong, Xiaofei Xie, Minhui Xue, and Yang Liu. Breaking Neural Reasoning Architectures with Metamorphic Relation-Based Adversarial Examples. *IEEE Transactions on Neural Networks and Learning Systems (TNNLS)*, 2021.

[3] Yupeng Cheng, Qing Guo, Felix Juefei-Xu, Xiaofei Xie, Shang-Wei Lin, Weisi Lin, Wei Feng, and Yang Liu. Pasadena: Perceptually Aware and Stealthy Adversarial Denoise Attack. *IEEE Transactions on Multimedia (TMM)*, 2021.

[4] Ziyi Cheng, Xuhong Ren, Felix Juefei-Xu, Wanli Xue, Qing Guo, Lei Ma, and Jianjun Zhao. DeepMix: Online Auto Data Augmentation for Robust Visual Object Tracking. In *IEEE International Conference on Multimedia and Expo (ICME)*, 2021.

[5] Lan Fu, Qing Guo, Felix Juefei-Xu, Hongkai Yu, Wei Feng, Yang Liu, and Song Wang. Benchmarking Shadow Removal for Facial Landmark Detection and Beyond. *arXiv preprint arXiv:2111.13790*, 2021.

[6] Lan Fu, Hongkai Yu, Felix Juefei-Xu, Jinlong Li, Qing Guo, and Song Wang. Let There be Light: Improved Traffic Surveillance via Detail Preserving Night-to-Day Transfer. *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*, 2021.

[7] Lan Fu, Changqing Zhou, Qing Guo, Felix Juefei-Xu, Hongkai Yu, Wei Feng, Yang Liu, and Song Wang. Auto-Exposure Fusion for Single-Image Shadow Removal. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, June 2021.

[8] Ruijun Gao, Qing Guo, Felix Juefei-Xu, Hongkai Yu, and Wei Feng. AdvHaze: Adversarial Haze Attack. *arXiv preprint arXiv:2104.13673*, 2021.

[9] Ruijun Gao, Qing Guo, Felix Juefei-Xu, Hongkai Yu, Huazhu Fu, Wei Feng, Yang Liu, and Song Wang. Can You Spot the Chameleon? Adversarially Camouflaging Images from Co-Salient Object Detection. *arXiv preprint arXiv:2009.09258*, 2021.

[10] Ruijun Gao, Qing Guo, Qian Zhang, Felix Juefei-Xu, Hongkai Yu, and Wei Feng. Adversarial Relighting against Face Recognition. *arXiv preprint arXiv:2108.07920*, 2021.

[11] Qing Guo, Ziyi Cheng, Felix Juefei-Xu, Lei Ma, Xiaofei Xie, Yang Liu, and Jianjun Zhao. Learning to Adversarially Blur Visual Object Tracking. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. IEEE, October 2021.

[12] Qing Guo, Felix Juefei-Xu, Xiaofei Xie, Lei Ma, Jian Wang, Bing Yu, Wei Feng, and Yang Liu. Watch out! Motion is Blurring the Vision of Your Deep Neural Networks. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

[13] Qing Guo, Felix Juefei-Xu, Changqing Zhou, Yang Liu, and Song Wang. Sparta: Spatially Attentive and Adversarially Robust Activation. *arXiv preprint arXiv:2105.08269*, 2021.

[14] Qing Guo, Xiaoguang Li, Felix Juefei-Xu, Hongkai Yu, Yang Liu, and Song Wang. JPGNet: Joint Predictive Filtering and Generative Network for Image Inpainting. In *Proceedings of the ACM International Conference on Multimedia (ACM MM)*, 2021.

[15] Qing Guo, Jingyang Sun, Felix Juefei-Xu, Lei Ma, Xiaofei Xie, Wei Feng, and Yang Liu. EfficientDeRain: Learning Pixel-wise Dilation Filtering for High-Efficiency Single-Image Deraining. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*. AAAI, Feb 2021.

[16] Qing Guo, Zhijie Wang, Felix Juefei-Xu, Di Lin, Lei Ma, Wei Feng, and Yang Liu. CarveNet: Carving Point-Block for Complex 3D Shape Completion. *arXiv preprint arXiv:2107.13452*, 2021.

[17] Qing Guo, Xiaofei Xie, Felix Juefei-Xu, Lei Ma, Zhongguo Li, Wanli Xue, Wei Feng, and Yang Liu. SPARK: Spatial-aware Online Incremental Attack Against Visual Tracking. In *Proceedings of the European Conference on Computer Vision (ECCV)*, Aug 2020.

[18] Yihao Huang, Qing Guo, Felix Juefei-Xu, Lei Ma, Weikai Miao, Yang Liu, and Geguang Pu. AdvFilter: Predictive Perturbation-aware Filtering against Adversarial Attack via Multi-domain Learning. In *Proceedings of the ACM International Conference on Multimedia (ACM MM)*, 2021.

[19] Yihao Huang, Felix Juefei-Xu, Qing Guo, Yang Liu, and Geguang Pu. FakeLocator: Robust Localization of GAN-Based Face Manipulations. *IEEE Transactions on Information Forensics and Security (TIFS)*, 2022.

[20] Yihao Huang, Felix Juefei-Xu, Qing Guo, Lei Ma, Xiaofei Xie, Weikai Miao, Yang Liu, and Geguang Pu. Dodging DeepFake Detection via Implicit Spatial-Domain Notch Filtering. *arXiv preprint arXiv:2009.09213*, 2021.

[21] Yihao Huang, Felix Juefei-Xu, Qing Guo, Weikai Miao, Yang Liu, and Geguang Pu. AdvBokeh: Learning to Adversarially Defocus Blur. *arXiv preprint arXiv:2111.12971*, 2021.

[22] Yihao Huang, Felix Juefei-Xu, Run Wang, Qing Guo, Lei Ma, Xiaofei Xie, Jianwen Li, Weikai Miao, Yang Liu, and Geguang Pu. FakePolisher: Making DeepFakes More Detection-Evasive by Shallow Reconstruction. In *Proceedings of the ACM International Conference on Multimedia (ACM MM)*, 2020.

[23] Felix Juefei-Xu, Vishnu Naresh Boddeti, and Marios Savvides. Local Binary Convolutional Neural Networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 19–28. IEEE, July 2017.

[24] Felix Juefei-Xu, Vishnu Naresh Boddeti, and Marios Savvides. Perturbative Neural Networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3310–3318. IEEE, June 2018.

[25] Felix Juefei-Xu, Rahul Dey, Vishnu Naresh Bodetti, and Marios Savvides. RankGAN: A Maximum Margin Ranking GAN for Generating Faces. In *Proceedings of the Asian Conference on Computer Vision (ACCV)*, December 2018.

[26] Felix Juefei-Xu, Run Wang, Yihao Huang, Qing Guo, Lei Ma, and Yang Liu. Countering Malicious DeepFakes: Survey, Battleground, and Horizon. *arXiv preprint arXiv:2103.00218*, 2021.

[27] Yiming Li, Congcong Wen, Felix Juefei-Xu, and Chen Feng. Fooling LiDAR Perception via Adversarial Trajectory Perturbation. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. IEEE, October 2021.

[28] Lei Ma, Felix Juefei-Xu, Jiyuan Sun, Chunyang Chen, Ting Su, Fuyuan Zhang, Minhui Xue, Bo Li, Li Li, Yang Liu, Jianjun Zhao, and Yadong Wang. DeepGauge: Multi-Granularity Testing Criteria for Deep Learning Systems. In *The 33rd IEEE/ACM International Conference on Automated Software Engineering (ASE)*, 2018.

[29] Lei Ma, Felix Juefei-Xu, Minhui Xue, Qiang Hu, Sen Chen, Bo Li, Yang Liu, Jianjun Zhao, Jianxiong Yin, and Simon See. Secure Deep Learning Engineering: A Software Quality Assurance Perspective. *arXiv preprint arXiv:1810.04538*, 2018.

[30] Lei Ma, Felix Juefei-Xu, Minhui Xue, Bo Li, Li Li, Yang Liu, and Jianjun Zhao. DeepCT: Tomographic Combinatorial Testing for Deep Learning Systems. *Proceedings of the IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)*, 2019.

[31] Lei Ma, Fuyuan Zhang, Jiyuan Sun, Minhui Xue, Bo Li, Felix Juefei-Xu, Chao Xie, Li Li, Yang Liu, Jianjun Zhao, and Yadong Wang. DeepMutation: Mutation Testing of Deep Learning Systems. In *The 29th IEEE International Symposium on Software Reliability Engineering (ISSRE)*, 2018.

[32] Hua Qi, Qing Guo, Felix Juefei-Xu, Xiaofei Xie, Lei Ma, Wei Feng, Yang Liu, and Jianjun Zhao. DeepRhythm: Exposing DeepFakes with Attentional Visual Heartbeat Rhythms. In *Proceedings of the ACM International Conference on Multimedia (ACM MM)*, 2020.

[33] Hua Qi, Zhijie Wang, Qing Guo, Jianlang Chen, Felix Juefei-Xu, Lei Ma, and Jianjun Zhao. ArchRepair: Block-Level Architecture-Oriented Repairing for Deep Neural Networks. *arXiv preprint arXiv:2111.13330*, 2021.

[34] Binyu Tian, Qing Guo, Felix Juefei-Xu, Wen Le Chan, Yupeng Cheng, Xiaohong Li, Xiaofei Xie, and Shengchao Qin. Bias Field Poses a Threat to DNN-Based X-Ray Recognition. In *IEEE International Conference on Multimedia and Expo (ICME)*, 2021.

[35] Binyu Tian, Felix Juefei-Xu, Qing Guo, Xiaofei Xie, Xiaohong Li, and Yang Liu. AVA: Adversarial Vignetting Attack against Visual Recognition. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2021.

[36] Run Wang, Felix Juefei-Xu, Qing Guo, Yihao Huang, Xiaofei Xie, Lei Ma, and Yang Liu. Amora: Black-box Adversarial Morphing Attack. In *Proceedings of the ACM International Conference on Multimedia (ACM MM)*, 2020.

[37] Run Wang, Felix Juefei-Xu, Yihao Huang, Qing Guo, Xiaofei Xie, Lei Ma, and Yang Liu. DeepSonar: Towards Effective and Robust Detection of AI-Synthesized Fake Voices. In *Proceedings of the ACM International Conference on Multimedia (ACM MM)*, 2020.

[38] Run Wang, Felix Juefei-Xu, Meng Luo, Yang Liu, and Lina Wang. JPGNet: Joint Predictive Filtering and Generative Network for Image Inpainting. In *Proceedings of the ACM International Conference on Multimedia (ACM MM)*, 2021.

[39] Run Wang, Felix Juefei-Xu, Lei Ma, Xiaofei Xie, Yihao Huang, Jian Wang, and Yang Liu. FakeSpotter: A Simple yet Robust Baseline for Spotting AI-Synthesized Fake Faces. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2020.

[40] Xiaofei Xie, Lei Ma, Felix Juefei-Xu, Minhui Xue, Hongxu Chen, Yang Liu, Jianjun Zhao, Bo Li, Jianxiong Yin, and Simon See. DeepHunter: A Coverage-Guided Fuzz Testing Framework for Deep Neural Networks. In *ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA)*, 2019.

[41] Bing Yu, Hua Qi, Qing Guo, Felix Juefei-Xu, Xiaofei Xie, Lei Ma, and Jianjun Zhao. DeepRepair: Style-Guided Repairing for DNNs in the Real-world Operational Environment. *IEEE Transactions on Reliability (TR)*, 2021.

[42] Liming Zhai, Felix Juefei-Xu, Qing Guo, Xiaofei Xie, Lei Ma, Wei Feng, Shengchao Qin, and Yang Liu. It's raining cats or dogs? adversarial rain attack on dnn perception. *arXiv preprint arXiv:2009.09205*, 2020.