1. 目标检测

[1] Girshick R, Donahue J, Darrell T, et al. Region-based convolutional networks for accurate object detection and segmentation[J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 38(1): 142-158.

[2] Girshick R. Fast r-cnn[C]//Proceedings of the IEEE international conference on computer vision. 2015: 1440-1448.

[3] Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks[J]. Advances in neural information processing systems, 2015, 28.

[4] He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 37(9): 1904-1916.

[5] He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 37(9): 1904-1916.

[6] He K, Gkioxari G, Dollár P, et al. Mask r-cnn[C]//Proceedings of the IEEE international conference on computer vision. 2017: 2961-2969.

[7]Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 779-788.

[8] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C]//European conference on computer vision. Springer, Cham, 2016: 21-37.

[9] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE international conference on computer vision. 2017: 2980-2988.

[10] Zhao Q, Sheng T, Wang Y, et al. M2det: A single-shot object detector based on multi-level feature pyramid network[C]//Proceedings of the AAAI conference on artificial intelligence. 2019, 33(01): 9259-9266.

1. 目标追踪

[1] Bewley A, Ge Z, Ott L, et al. Simple online and realtime tracking[C]//2016 IEEE international conference on image processing (ICIP). IEEE, 2016: 3464-3468.

[2] Wojke N, Bewley A, Paulus D. Simple online and realtime tracking with a deep association metric[C]//2017 IEEE international conference on image processing (ICIP). IEEE, 2017: 3645-3649.

[3] Voigtlaender P, Krause M, Osep A, et al. Mots: Multi-object tracking and segmentation[C]//Proceedings of the ieee/cvf conference on computer vision and pattern recognition. 2019: 7942-7951.

[4] Wang Z, Zheng L, Liu Y, et al. Towards real-time multi-object tracking[C]//European Conference on Computer Vision. Springer, Cham, 2020: 107-122.

[5] Zhou X, Koltun V, Krähenbühl P. Tracking objects as points[C]//European Conference on Computer Vision. Springer, Cham, 2020: 474-490.

[6] Zhang Y, Wang C, Wang X, et al. Fairmot: On the fairness of detection and re-identification in multiple object tracking[J]. International Journal of Computer Vision, 2021, 129(11): 3069-3087.

[7] Xu Z, Zhang W, Tan X, et al. Segment as points for efficient online multi-object tracking and segmentation[C]//European Conference on Computer Vision. Springer, Cham, 2020: 264-281.

[8] Sun P, Cao J, Jiang Y, et al. Transtrack: Multiple object tracking with transformer[J]. arXiv preprint arXiv:2012.15460, 2020.

[9] Meinhardt T, Kirillov A, Leal-Taixe L, et al. Trackformer: Multi-object tracking with transformers[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 8844-8854.

[10] Chu P, Wang J, You Q, et al. Transmot: Spatial-temporal graph transformer for multiple object tracking[J]. arXiv preprint arXiv:2104.00194, 2021.

1. 对抗攻击

[1] Goodfellow I J, Shlens J, Szegedy C. Explaining and harnessing adversarial examples[J]. arXiv preprint arXiv:1412.6572, 2014.

[2] Kurakin A, Goodfellow I, Bengio S. Adversarial machine learning at scale[J]. arXiv preprint arXiv:1611.01236, 2016.

[3] Dong Y, Liao F, Pang T, et al. Boosting adversarial attacks with momentum[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 9185-9193.

[4] Dong Y, Pang T, Su H, et al. Evading defenses to transferable adversarial examples by translation-invariant attacks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 4312-4321.

[5] Xie C, Zhang Z, Zhou Y, et al. Improving transferability of adversarial examples with input diversity[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 2730-2739.

[6] Li M, Deng C, Li T, et al. Towards transferable targeted attack[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 641-649.

[7] Zou J, Pan Z, Qiu J, et al. Improving the transferability of adversarial examples with resized-diverse-inputs, diversity-ensemble and region fitting[C]//European Conference on Computer Vision. Springer, Cham, 2020: 563-579.

[8] Gao L, Zhang Q, Song J, et al. Patch-wise attack for fooling deep neural network[C]//European Conference on Computer Vision. Springer, Cham, 2020: 307-322.

[9] Lin J, Song C, He K, et al. Nesterov accelerated gradient and scale invariance for adversarial attacks[J]. arXiv preprint arXiv:1908.06281, 2019.

[10] Wu L, Zhu Z, Tai C. Understanding and enhancing the transferability of adversarial examples[J]. arXiv preprint arXiv:1802.09707, 2018.

1. Transformer相关

[1] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.

[2] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J]. arXiv preprint arXiv:2010.11929, 2020.

[3] Liu Z, Lin Y, Cao Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 10012-10022.

[4] Mehta S, Rastegari M. Mobilevit: light-weight, general-purpose, and mobile-friendly vision transformer[J]. arXiv preprint arXiv:2110.02178, 2021.

[5] Pan J, Bulat A, Tan F, et al. EdgeViTs: Competing Light-weight CNNs on Mobile Devices with Vision Transformers[J]. arXiv preprint arXiv:2205.03436, 2022.

[6] Wang W, Xie E, Li X, et al. Pyramid vision transformer: A versatile backbone for dense prediction without convolutions[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 568-578.

1. 弱监督学习

[1] Bilen H, Vedaldi A. Weakly supervised deep detection networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 2846-2854.

[2] Tang P, Wang X, Bai X, et al. Multiple instance detection network with online instance classifier refinement[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2843-2851.

[3] Tang P, Wang X, Bai S, et al. Pcl: Proposal cluster learning for weakly supervised object detection[J]. IEEE transactions on pattern analysis and machine intelligence, 2018, 42(1): 176-191.

[4] Jiang P T, Hou Q, Cao Y, et al. Integral object mining via online attention accumulation[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2019: 2070-2079.

[5] Lin C, Wang S, Xu D, et al. Object instance mining for weakly supervised object detection[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2020, 34(07): 11482-11489.

[6] Ren Z, Yu Z, Yang X, et al. Instance-aware, context-focused, and memory-efficient weakly supervised object detection[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 10598-10607.

[7] Chen Z, Fu Z, Jiang R, et al. Slv: Spatial likelihood voting for weakly supervised object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 12995-13004.

[8] Shen Y, Ji R, Chen Z, et al. UWSOD: Toward fully-supervised-level capacity weakly supervised object detection[J]. Advances in Neural Information Processing Systems, 2020, 33: 7005-7019.

[9] Cheng G, Yang J, Gao D, et al. High-quality proposals for weakly supervised object detection[J]. IEEE Transactions on Image Processing, 2020, 29: 5794-5804.

[10] Xu Y, Zhou C, Yu X, et al. Pyramidal multiple instance detection network with mask guided self-correction for weakly supervised object detection[J]. IEEE Transactions on Image Processing, 2021, 30: 3029-3040.