Wenhai Wang

Email: wangwenhai[at]pjlab.org.cn wangwenhai362[at]163.com

Homepage: https://whai362.github.io

Education

Nanjing University, Nanjing, China Sep. 2016 - Sep. 2021

Ph.D. in Computer Science and Technology, *Supervised by Prof. Tong Lu*.

Nanjing University of Science and Technology, Nanjing, China Sep. 2012 - Jun. 2016

B.E. in Software Engineering.

Research Interests

Large-Scale Vision/Vision-Language Foundation Backbone
Object Detection & Segmentation
Autonomous Driving Perception
Optical Character Recognition

Experience

Github: https://github.com/whai362

Shanghai AI Laboratory, Shanghai, China Sep. 2021 - Present

Research Scientist, *Collaborated with Prof. Jifeng Dai and Prof. Yu Qiao*
Leading a team and working on (1) large-scale vision/vision-language foundation model, (2) vision Transformer, and (3) autonomous driving perception.

The University of Hong Kong, Hongkong, China Oct. 2019 - Mar. 2020

Research Assistant, *Supervised by Prof. Ping Luo*
Worked on instance segmentation, and optical character recognition.

SenseTime, Beijing, China Aug. 2019 - Mar. 2020

Research Intern, *Supervised by Xuebo Liu and Ding Liang*
Worked on optical character recognition.

Momenta, Beijing, China Jun. 2018 - Dec. 2018

Research Intern, *Supervised by Prof. Xiang Li*
Worked on CNN backbone, and object detection.

Honors and Awards

Waymo 2022 3D Camera-Only Detection Task, **1st Place, 15,000 USD Bonus** June 2022

National Artificial Intelligence Challenge (NAIC) 2020 , Remote Sensing Semantic Segmentation Task, **1st Place, 1,000,000 RMB Bonus** Dec. 2020

China National Scholarship for Doctoral Students Dec. 2019

ICDAR2019 Robust Reading Challenge on Arbitrary-Shaped Text, Task1, **1st Place** Sept. 2019

ICDAR2019 Robust Reading Challenge on Large-scale Street View Text with Partial Labeling, Task1, **2nd Place** Sept. 2019

AI Challenger 2018 Autonomous Driving Perception Task, **2nd Place, 40,000 RMB Bonus** Dec. 2018

ACM-ICPC Asia Regional Contest, **Silver Medal** Nov. 2015

PUBLICATIONS

Google Scholar: https://scholar.google.com/citations?user=WM0OglcAAAAJ
(* Equal contribution, † Interns, # Corresponding authors)

[J1] **W. Wang***, E. Xie*, X. Li, et al. PAN++: Towards Efficient and Accurate End-to-End Spotting of Arbitrarily-Shaped Text[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI). 2021.

[J2] **W. Wang#**, E. Xie, X. Li, et al. PVT v2: Improved Baselines with Pyramid Vision Transformer[J]. Computational Visual Media Journal (CVMJ). 2022. (ESI Highly Cited Paper (1%), Honorable Mention Award)

[J3] **W. Wang**, Z. Li, T. Lu#. Grid Dividing for Single-Stage Instance Segmentation[J]. Journal of Software (JoS), 2021 (in Chinese).

[J4] E. Xie*, **W. Wang***, M. Ding, et al. PolarMask++: Enhanced Polar Representation for Single-Shot Instance Segmentation and Beyond[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI). 2021.

[J5] B. Dong, **W. Wang**, D. Fan#, et al. Polyp-PVT: Polyp Segmentation with Pyramid Vision Transformers[J]. CAAI Artificial Intelligence Research (CAAI AIR). 2021.

[J6] X. Li, C. Lv, **W. Wang**, et al. Generalized Focal Loss: Towards Efficient Representation Learning for Dense Object Detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI). 2022.

[J7] R. Liu, Z. Pang, Z. Meng, **W. Wang**, et al. On efficient reinforcement learning for full-length game of StarCraft II[J]. Journal of Artificial Intelligence Research (JAIR). 2022.

[C8] **W. Wang***, J. Dai*, Z. Chen*, et al. InternImage: Exploring Large-Scale Vision Foundation Models with Deformable Convolutions[C] // Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2023. (Highlight Paper (2.5%))

[C9] **W. Wang**, E. Xie, X. Li, et al. Pyramid Vision Transformer: A Versatile Backbone for Dense Prediction without Convolutions[C] // Proceedings of IEEE/CVF International Conference on Computer Vision (ICCV). 2021. (Oral Presentation (3.4%)) (ICCV21' Top-10 Influential Papers (Rank 2))

[C10] **W. Wang**, X. Liu, X. Ji, et al. AE TextSpotter: Learning Visual and Linguistic Representation for Ambiguous Text Spotting[C] // Proceedings of the European Conference on Computer Vision (ECCV). 2020.

[C11] **W. Wang***, E. Xie*, X. Song, et al. Efficient and Accurate Arbitrary-Shaped Text Detection with Pixel Aggregation Network[C] // Proceedings of IEEE/CVF International Conference on Computer Vision (ICCV). 2019.

[C12] **W. Wang***, E. Xie*, X. Li, et al. Shape Robust Text Detection with Progressive Scale Expansion Network[C] // Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019.

[C13] **W. Wang***, X. Li*, T. Lu#, et al. Mixed Link Networks[C] // Proceedings of International Joint Conference on Artificial Intelligence (IJCAI). 2018. (Oral Presentation)

[C14] Z. Chen†, Y. Duan†, **W. Wang#**, et al. ViT-Adapter: Exploring Plain Vision Transformer for Accurate Dense Predictions[C] // Proceedings of the the International Conference on Learning Representations (ICLR). 2023. (Spotlight Paper (8.0%))

[C15] Z. Li*†, **W. Wang***, H. Li*, et al. BEVFormer: Learning Bird's-Eye-View Representation from Multi-Camera Images via Spatiotemporal Transformers[C] // Proceedings of the European Conference on Computer Vision (ECCV). 2022.

[C16] C. Tian*†, **W. Wang***, X. Zhu, et al. VL-LTR: Learning Class-wise Visual-Linguistic Representation for Long-Tailed Visual Recognition[C] // Proceedings of the European Conference on Computer Vision (ECCV). 2022.

[C17] Z. Li†, **W. Wang#**, E. Xie, et al. Panoptic SegFormer: Delving Deeper into Panoptic Segmentation with Transformers[C] // Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022.

[C18] Z. Chen†, **W. Wang#**, E. Xie, et al. Towards Ultra-Resolution Neural Style Transfer via Thumbnail Instance Normalization[C] // Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI). 2022.

[C19] E. Xie, **W. Wang**, Z. Yu, et al. SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers[C] // Advances in Neural Information Processing Systems (NeurIPS). 2021. (NeurIPS21' Top-10 Influential Papers (Rank 3))

[C20] X. Li, **W. Wang**, X. Hu, et al. Generalized Focal Loss V2: Learning Reliable Localization Quality Esti-

mation for Dense Object Detection[C] // Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2021.

[C21] X. Li, **W. Wang**, L. Wu, et al. Generalized Focal Loss: Learning Qualified and Distributed Bounding Boxes for Dense Object Detection[C] // Advances in Neural Information Processing Systems (NeurIPS). 2020.

[C22] X. Li, **W. Wang**, X. Hu, et al. Selective Kernel Networks[C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019.

[C23] J. Zhu, X. Zhu, **W. Wang**, et al. Uni-Perceiver-MoE: Learning Sparse Generalist Models with Conditional MoEs[C] // Advances in Neural Information Processing Systems (NeurIPS). 2022.

[C24] E. Xie, J. Ding, **W. Wang**, et al. Detco: Unsupervised contrastive learning for object detection[C] // Proceedings of IEEE/CVF International Conference on Computer Vision (ICCV). 2021.

[C25] E. Xie, W. Wang, **W. Wang**, et al. Segmenting Transparent Objects in the Wild[C] // Proceedings of the European Conference on Computer Vision (ECCV). 2020.

[C26] E. Xie, W. Wang, **W. Wang**, et al. Segmenting Transparent Object in the Wild with Transformer[C] // Proceedings of International Joint Conference on Artificial Intelligence (IJCAI). 2021.

[C27] E. Xie, P. Sun, X. Song, **W. Wang**, et al. PolarMask: Single Shot Instance Segmentation with Polar Representation[C] // Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020. (Oral Presentation (5.7%)) (CVPR20' Top-10 Influential Papers)

[C28] S. Jin, W. Liu, E Xie, **W. Wang**, et al. Differentiable Hierarchical Graph Grouping for Multi-Person Pose Estimation[C] // Proceedings of the European Conference on Computer Vision (ECCV). 2020.

[C29] W. Wang, E. Xie, X. Liu, **W. Wang**, et al. Scene Text Image Super-Resolution in the Wild[C] // Proceedings of the European Conference on Computer Vision (ECCV). 2020.

[C30] Y. Hu, J. Yang, L. Chen, ..., **W. Wang**, et al. Goal-oriented Autonomous Driving[C] // Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2023. (Highlight Paper (2.5%))

[C31] H. Li, J. Zhu, X. Jiang, ..., **W. Wang**, et al. Uni-Perceiver v2: A Generalist Model for Large-Scale Vision and Vision-Language Tasks[C] // Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2023. (Award Candidate (12 out of 9155)))

[C32] G. Shi, Y. Wu#, J. Liu, S. Wan, **W. Wang**, et al. Incremental Few-Shot Semantic Segmentation via Embedding Adaptive-Update and Hyper-class Representation[C] // Proceedings of the 30th ACM International Conference on Multimedia (ACM MM). 2022.

Academic Service

Senior Program Committee Member

- International Joint Conference on Artificial Intelligence (IJCAI), 2021

Journal Reviewer

- IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)
- International Journal of Computer Vision (IJCV)
- IEEE Transactions on Image Processing (TIP)
- IEEE Transactions on Multimedia (TMM)
- Computational Visual Media Journal (CVMJ)
- Pattern Recognition (PR)

Program Committee Member/Conference Reviewer

- IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, 2021, 2022, 2023
- Neural Information Processing Systems (NeurIPS), 2020, 2021, 2023
- International Conference on Machine Learning (ICML), 2021, 2022
- International Conference on Learning Representations (ICLR), 2021
- IEEE/CVF International Conference on Computer Vision (ICCV), 2021
- European Conference on Computer Vision (ECCV), 2022
- AAAI Conference on Artificial Intelligence (AAAI), 2022
- International Joint Conference on Artificial Intelligence (IJCAI), 2022
- Asian Conference on Computer Vision 2020 (ACCV), 2020
- IEEE Winter Conference on Applications of Computer Vision (WACV), 2021