Wenhai Wang

Email: wangwenhai362[at]{163.com, smail.nju.edu.cn} wangwenhai[at]pjlab.org.cn

Homepage: https://whai362.github.io

Education

Nanjing University, Nanjing, China Sep. 2018 - Sep. 2021

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

Nanjing University, Nanjing, China Sep. 2016 - Jun. 2018

M.S. 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

- Vision-Language Model
- CNN / Transformer Backbone
- Object Detection & Semantic/Instance/Panoptic Segmentation
- Scene Text Detection & Recognition
- Autonomous Driving

Experience

Github: https://github.com/whai362

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

Young Scientist, *Supervised by Dr. Jifeng Dai and Prof. Yu Qiao*

- **Vision-Language Model**. Improving vision-language models and applying vision-language models to downstream recognition tasks.
- **Vision Transformer**. Designing effective and efficient vision transformer encoder / decoder for image recognition tasks, such as classification, detection and segmentation.
- **Autonomous Driving**. Exploring generic visual representations for various perception tasks in the autonomous driving scenario.

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

Research Assistant, *Supervised by Prof. Ping Luo*

- **One-stage Instance Segmentation**. Proposed an anchor-box free and single shot instance segmentation method, termed PolarMask, which formulates the instance segmentation problem as instance center classification and dense distance regression in a polar coordinate. This work *PolarMask: Single Shot Instance Segmentation with Polar Representation* is accepted by CVPR 2020, and the extension version is accepted by TPAMI 2021.

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

Research Intern, *Supervised by Xuebo Liu and Ding Liang*

- **Optical Character Recognition**. Did research on text detection and recognition in natural scenes, and tried to 1) solve the ambiguity in text detection and 2) design a real-time algorithm for arbitrary-shaped text detection and recognition. The first work is accepted by ECCV 2020, and the second work is accepted by TPAMI 2021.

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

Research Intern, *Supervised by Dr. Xiang Li*

- **Deep Neural Networks Exploration**. Did research on the receptive fields of CNNs, and designed a dynamic selection mechanism in CNNs that allows each neuron to adaptively adjust its receptive field. This work *Selective Kernel Networks* is accepted by CVPR 2019.
- **Object Detection**. Led a team to take part in Autonomous Driving Perception Task in AI Challenger 2018. Re-implemented and improved Cascade R-CNN to detect objections in autonomous driving scenarios. Our team finally secured the 2nd place in the task.

CONTESTS

- National Artificial Intelligence Challenge (NAIC) 2020 , Remote Sensing Semantic Segmentation Task, **1st Place, 1,000,000 RMB** Nov. 2020
- ICDAR2019 Robust Reading Challenge on Arbitrary-Shaped Text, Task1, **1st Place** May 2019
- ICDAR2019 Robust Reading Challenge on Large-scale Street View Text with Partial Labeling, Task1, **2nd Place** Jun. 2019
- AI Challenger 2018 Autonomous Driving Perception Task, **2nd Place, 40,000 RMB** Dec. 2018
- ACM-ICPC Asia Regional Contest, **Silver Medal** Nov. 2015

HONORS AND AWARDS

- China National Scholarship (the highest scholarship for students studying in China) Oct. 2019

PUBLICATIONS

Google Scholar: `https://scholar.google.com/citations?user=WM0OglcAAAAJ`

Top-Tier Computer Vision Journal & Conference

(* indicates equal contribution, # corresponding author)

[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] 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.

[C1] **W. Wang**, E. Xie, X. Li, et al. Pyramid Vision Transformer: A Versatile Backbone for Dense Prediction without Convolutions[C] // Proceedings of IEEE International Conference on Computer Vision (ICCV). 2021. (Oral)

[C2] **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.

[C3] **W. Wang***, E. Xie*, X. Song, et al. Efficient and Accurate Arbitrary-Shaped Text Detection with Pixel Aggregation Network[C] // Proceedings of IEEE International Conference on Computer Vision (ICCV). 2019.

[C4] **W. Wang***, E. Xie*, X. Li, et al. Shape Robust Text Detection with Progressive Scale Expansion Network[C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2019.

[C5] **W. Wang***, X. Li*, T. Lu#, et al. Mixed Link Networks[C] // Proceedings of International Joint Conference on Artificial Intelligence (IJCAI). 2018. (Oral)

[C6] 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.

[C7] 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.

[C8] X. Li, **W. Wang**, X. Hu, et al. Generalized Focal Loss V2: Learning Reliable Localization Quality Estimation for Dense Object Detection[C] // Proceedings of IEEE Conference on Computer Vision and Pattern

Recognition (CVPR). 2021.

[C9] 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.

[C10] X. Li, **W. Wang**, X. Hu, et al. Selective Kernel Networks[C] // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2019.

[C11] E. Xie, J. Ding, **W. Wang**, et al. Detco: Unsupervised contrastive learning for object detection[C] // Proceedings of IEEE International Conference on Computer Vision (ICCV). 2021.

[C12] 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.

[C13] 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.

[C14] E. Xie, P. Sun, X. Song, **W. Wang**, et al. PolarMask: Single Shot Instance Segmentation with Polar Representation[C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2020. (Oral)

[C15] 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.

[C16] 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.

Academic Service

Journal Reviewer

- IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)
- IEEE Transactions on Image Processing (TIP)
- IEEE Transactions on Multimedia (TMM)
- Computational Visual Media Journal (CVMJ)

(Senior) Program Committee Member/Conference Reviewer

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