JIANQI MA

Tel: +86 18817362168 | Mail: mjq11302010044@gmail.com | Github: https://github.com/mjq11302010044/ (Earned 1K+ stars)

Google Scholar: https://scholar.google.com/citations?user=kQUJjQQAAAAJ&hl=en (1K+ citations)

Research Interests: OCR, Detection, Segmentation, Image Super-resolution, Diffusion Models

EDUCATION

Doctor of Philosophy, Computing — Sep 2020 – Now

The Hong Kong Polytechnic University, Hong Kong, China

Master of Science, Computer Science — Sep 2016 – Jun 2019

Fudan University, Shanghai China

Bachelor of Software Engineering, — Sep 2011 – Jun 2015

Fudan University, Shanghai China

PUBLICATIONS

Scene Text Image Super-resolution:

[1] Ma, J., Liang, Z., & Zhang, L. (2022). A Text Attention Network for Spatial Deformation Robust Scene Text Image Super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 5911-5920. 2022.

[2] Ma, J., Guo, S., & Zhang, L. (2023). Text Prior Guided Scene Text Image Super-resolution. IEEE Transactions on Image Processing 32 (2023): 1341-1353

[3] Ma, J., Liang, Z., Xiang, W., Yang, X., & Zhang, L. (2023). A Benchmark for Chinese-English Scene Text Image Super-resolution. (Recently accepted by ICCV 2023)

[4] Chen, J., Yu, H., **Ma, J.,** Li, B., & Xue, X. (2022). Text gestalt: Stroke-aware scene text image super-resolution. In Proceedings of the AAAI Conference on Artificial Intelligence, vol. 36, no. 1, pp. 285-293.

Scene Text Detection and Segmentation:

[5] Ma, J., Shao, W., Ye, H., Wang, L., Wang, H., Zheng, Y., & Xue, X. (2018). Arbitrary-oriented scene text detection via rotation proposals. IEEE Transactions on Multimedia, 20(11), 3111-3122. **(Citation: 1100+)**

[6] Ma, J. (2020). RRPN++: Guidance towards more accurate scene text detection. arXiv preprint arXiv:2009.13118.

[7] Xu, X., Qi, Z., **Ma, J.,** Zhang, H., Shan, Y., & Qie, X. (2022). BTS: A Bi-lingual Benchmark for Text Segmentation in the Wild. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 19152-19162. **(Oral paper)**

Scene Text Recognition:

[8] Yu, H., Chen, J., Li, B., **Ma, J.**, Guan, M., Xu, X., ... & Xue, X. (2021). Benchmarking chinese text recognition: Datasets, baselines, and an empirical study. arXiv preprint arXiv:2112.15093.

[8] Chinese Book Publication in OCR technique practice:
刘树春, 贺盼, **马建奇(Jianqi Ma)**, 王佳军, 谢雨飞, 陈明曦**《深度实践 OCR：基于深度学习的文字识别》**，机械工业出版社, 2020.6

Other Low-level Vision:

[9] Guo, S., Yang, X., **Ma, J.**, Ren, G., & Zhang, L. (2022). A differentiable two-stage alignment scheme for burst image reconstruction with large shift. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. , pp. 17472-17481.

[10] Guo, S., Yong, H., Zhang, X., **Ma, J.**, & Zhang, L. (2023). Spatial-Frequency Attention for Image Denoising. arXiv preprint arXiv:2302.13598.

WORK EXPERIENCE

Research Assistant — Mar 2016 – Jun 2017

Shanghai Advanced Research Institute, CAS

Mentor: Dr. Yingbin Zheng

- We focused on task of **irregular-oriented scene text detection**, and built a new

framework to adapt arbitrary-oriented text line named *RRPN*. Article **[5]** published on *Transaction on Multimedia* received over **1100** citations (Google scholar). It has been ported to **detectron2** project by **Facebook Inc**.

- We built a face recognition software system, which jointed video camera stream and face recognition together to perform a real-time face recognition system. My parts of work were to extract 4096-d feature from a facial image using CNN and joint the video stream with the face recognition model using Flask.

Research Intern Aug 2018 – Nov 2018

Detection Team, Beijing Institute of Megvii Technology Inc

Mentor: Boxun Li, Dr. Gang Yu

- **Investigated better detector model** suitable for ARM-based board. Made better use of the computational utility of the board and got better facial detection performance (5% improvement on AP) on the task. (RetinaNet-based detector)

Research Intern Jul 2021 – Nov 2021

DAMO Academy, Alibaba Inc,

Mentor: Gaofeng Ren, Lei Zhang

- Improving Scene Text Image Super-resolution (STISR) with CNN-Transformer model. Paper **[1]** was accepted to CVPR 2022.

Research Intern Jul 2022 – Now

OPPO Institude,

Mentor: Lei Zhang

- Constructing a benchmark for scene text image super-resolution for Chinese-English texts. Collecting images with iPhone series and pairing the images with different focal lens. Proposing new evaluation protocols for more practical scenes.
- Using diffusion models for scene text image super-resolution. Suppressing the generation variety for diffusion model and taming the diffusion model for fidelity-based text image restoration.

PROJECTS

(1) Constructing dataset for multi-lingual (mainly Chinese and English) scene text restoration and investigating approaches for Chinese scene text restoration, especially using diffusion model. (2) Exploring diffusion generative model in STISR task. — Mar 2023

Introducing text prior into scene text image super resolution (STISR) **[2]** model to help enhance the high-resolution text image generation. We propose TPGSR (Text Prior Guided Super Resolution) to boost the STISR performance by **3%~10% recognition accuracy** in STISR benchmark TextZoom (accepted to IEEE TIP). — Mar 2021

Upgrading RRPN to RRPN++ **[4]**, we use anchor-free FPN to replace the Rotation-Anchor in the first stage of the model. Adopting RRoI-Align and recognition head to boost the detection performance in ICDAR2015 (**89.5%** in F-measure with single scale testing). — Jun 2020

On ICDAR2015 focusing and incidental detection benchmark, we both ranked **1st place** and was 3% ahead of 2nd place by proposing RRPN models which was able to detect multi-oriented text line with tighter rotated bounding box. The paper was accepted to IEEE TMM with citation 1100+ (June, 2023). — Jan 2017

SKILLS

(Proficient)Python; Pytorch, Tensorflow, Caffe; Numpy tensor operation;