

Xiaomeng Yang

(+86)15356603660 | xiaomeng.rachel@gmail.com | [xiaomeng-yang.github.io](https://github.com/xiaomeng-yang)

EDUCATION

University of Chinese Academy of Sciences (UCAS)

M.S. Computer Science

Beijing, China

Sep. 2021 – Jun. 2024 (Expected)

Zhejiang University (ZJU)

B.E. Electrical and Computer Engineering

Zhejiang, China

Sep. 2017 – Jun. 2021

University of Illinois at Urbana-Champaign (UIUC)

B.S. Computer Engineering with Honors

Champaign, IL

Sep. 2017 – May 2021

PUBLICATIONS AND MANUSCRIPTS

- PIMNet++: A Parallel, Iterative and Diffusion-Based Network for Scene Text Recognition.
- **Xiaomeng Yang**, Zhi Qiao, Yu Zhou, Jin Wei, Yuan Ye, Zhilong Ji, Dongbao Yang, Weiping Wang. Masked and Permuted Implicit Context Learning for Scene Text Recognition. Under Review, 2023.
- **Xiaomeng Yang**, Dongbao Yang, Zhi Qiao, Yu Zhou. Accurate and Robust Scene Text Recognition via Adversarial Training. Under Review, 2023.
- Gangyan Zeng, Yuan Zhang, Yu Zhou, **Xiaomeng Yang**, Ning Jiang, Guoqing Zhao, Weiping Wang, Xu-Cheng Yin. Beyond OCR+VQA: Towards End-to-End Reading and Reasoning for Robust and Accurate TextVQA. In Pattern Recognition, Volume 138, 2023.
- Gangyan Zeng, Yuan Zhang, Yu Zhou, **Xiaomeng Yang**. Beyond OCR+VQA: Involving OCR into the Flow for Robust and Accurate TextVQA. In ACM MM, 2021. (Oral Presentation)
- Gangyan Zeng, Yuan Zhang, Yu Zhou, **Xiaomeng Yang**. A Cost-Efficient Framework for Scene Text Detection in the Wild. In PRICAI, 2021.
- Yucheng Jin, **Xiaomeng Yang**, Chengting Yu, Liangjing Yang. Educational Data Mining: Discovering Principal Factors for Better Academic Performance. In BDET, 2021.

RESEARCH EXPERIENCE

Linguistic Information Enhanced Scene Text Recognition

Jun. 2022 – Present

UCAS | Research Assistant, Advisor: Prof. Yu Zhou

Tomorrow Advancing Life | Research Internship, Advisor: Zhi Qiao

- Explored the relationship between the visual features and text information in scene text recognition.
- Proposed a framework for autoregressive decoding using masked, permuted implicit context learning and length supervision. Enhanced model robustness against length prediction errors through perturbation training. Our approach outperformed on challenging benchmarks. This work resulted in a top-tier conference submission.
- Explored non-autoregressive decoding in scene text recognition as image-based conditional text generation. Utilized discrete diffusion with a small transformer decoder and easy-first decoding for enhanced bidirectional context awareness. The manuscript is completed for submission to a top-tier journal.

Involving OCR into the Flow of TextVQA

Oct. 2020 – Sep. 2022

UCAS | Research Assistant, Advisor: Prof. Yu Zhou

- Developed two accurate TextVQA models by integrating OCR into the process.
- In the first scheme, designed a reading modification module for adaptive answer selection based on context. Visually enhanced text embedding and semantically oriented object embedding were used to enhance the feature representations. This approach, more robust to the OCR errors, led to a paper accepted at ACM MM 2021.
- In the second scheme, devised an efficient end-to-end network integrating text reading and reasoning, optimized by the downstream VQA signal. The approach outperformed existing alternatives in accuracy and robustness. This research led to a publication in Pattern Recognition 2023.

Adversarial Training for Scene Text Recognition

Apr. 2021 – Aug. 2021

UCAS | Summer Internship, Advisor: Prof. Yu Zhou

- Investigated the impact of adversarial training on scene text recognition.

- Implemented PGD attack on classical CTC and Attention-based methods, generating examples which is utilized for adversarial training. The model gained robustness and improved performance on low-resolution images with adversarial training. This work led to a short paper for ICASSP 2024 submission.

Cost-Efficient Scene Text Detection

Mar. 2021 – Jun. 2021

UCAS | *Summer Internship, Advisor: Prof. Yu Zhou*

- Explored unsupervised domain adaptation and active learning for cost-efficient data utilization.
- Proposed a two-stage framework using unsupervised domain adaptation with global transfer and text region transfer for pretraining, then fine-tuning with minimal actively annotated and pseudo-labeled real samples. The framework significantly outperforms the baseline. This work resulted in a paper accepted at PRICAI 2021.

Educational Data Mining

Oct. 2019 – Aug. 2020

ZJU | *SRTP Program, Advisor: Prof. Liangjing Yang*

- Studied educational data mining methods, which aims to discover knowledge from education-oriented databases.
- Utilized linear regression, regression tree, random forest, and neural network to conduct an investigation into the principal factors that influence Chinese junior high school students' academic performance on the China Education Panel Survey dataset. This work resulted in a paper accepted at BDET 2021.

Waterway Segmentation and Transfer Learning of Building Segmentation

Sep. 2019 – Dec. 2019

UIUC | *Individual Study, Advisor: Prof. Volodymyr Kindratenko*

- Started with the baseline segmentation model for straightforward building segmentation.
- Optimized the UNet model on specific DEM datasets, adjusting settings like overlapping rate and derivatives, and integrated a new Lovasz loss function for building and waterway segmentation, which enhances the baseline accuracy. Completed the course report with an 'A' grade.

ENGINEERING PROJECTS

Customs Express Information Extraction System

Aug. 2023 – Present

- Aiming to develop an information extraction system capable of processing diverse languages and formats from various countries and regions.
- The system initially identifies key information areas, like serial numbers, then detects detailed text sections. It recognizes texts in multiple languages and subsequently employs a visual information extraction model to parse and retrieve essential details from the text recognition results.
- Deployed and launched an initial version of the system.

Seal Recognition and General Character Recognition System

Jul. 2022 – Apr. 2023

- Created a universal character recognition system for financial documents, incorporating seal recognition.
- Investigated substituting the CNN backbone with ViT for Chinese document text recognition, maintaining comparable processing speed and facilitating handwritten text recognition. Tailored it for financial business applications.
- Delivered a universal character recognition system utilizing a ViT backbone and CTC decoder.

Scene Text Detection and Recognition Engine

Jan. 2022 – Oct. 2022

- Developed a scene text recognition engine for the Chinese National Radio and Television Administration.
- Enhanced the CTC-based recognition network with an image rectification preprocessing module and NLP post-processing for initial results refinement.
- The engine provided Traditional, Simplified Chinese, and English recognition, achieving about 95% accuracy in practical use.

SKILLS

Language skill: TOEFL 103 (Speaking 22)

Computer skills: Programming languages: Python, Java, , C/C++, JavaScript, HTML/CSS

Other Library and Tools: Numpy, Scipy, Cython, PyTorch, Tensroflow, Linux, Git

Libraries: NumPy, Matplotlib