

Yu-Ming Chen

Ph.D. Student in Computer Science and Technology @ Nankai University
Visual-Language Model & Agentic Workflow

<http://www.fishworld.site/>

EDUCATION

Nankai University (M.S. & Ph.D)

Computer Vision, supervised by Prof. Ming-Ming Cheng & Prof. Qi-Bin Hou

Tianjin, CN

Aug. 2022 – Present

Lanzhou University (B.S.)

Computer Science and Technology

Lanzhou, CN

Aug. 2018 – July 2022

PUBLICATIONS

Re-Aligning Language to Visual Objects with an Agentic Workflow

[Homepage]

Yu-Ming Chen, Jiang-Yan Feng, *et al.*

ICLR 2025

Agentic workflow including planning, tool use, and reflection steps to enhance LOD model performance by improving the alignment quality of language expressions and visual objects (data-centric).

YOLO-MS: rethinking multi-scale representation learning for real-time object detection

[Code]

Yu-Ming Chen, Xin-Bin Yuan, *et al.*

TPAMI 2025

By rethinking **multi-scale encoding** in multi-branch blocks, plug-and-play strategies are proposed to boost real-time detection accuracy with minimal computational overhead.

CrossKD: Cross-head knowledge distillation for object detection

[Code]

Jia-Bao Wang*, Yu-Ming Chen*, *et al.*

CVPR 2024

Relieving the student's head from receiving **contradictory supervision** signals from the ground-truth annotations and the teacher's predictions, improving the prediction mimicking to surpass the feature imitation.

Zone Evaluation: Revealing Spatial Bias in Object Detection

[Code]

Zhao-Hui Zheng, Yu-Ming Chen, *et al.*

TPAMI 2024

Identifies **spatial bias** in object detectors, revealing lower accuracy near borders, quantified by Zone Precision and linked to training data imbalance.

TeMO: Towards Text-Driven 3D Stylization for Multi-Object Meshes

[Code]

Xu-Ying Zhang, Yin-Bo Wen, Yu-Ming Chen *et al.*

CVPR 2024

Advances text-driven stylization for **multi-object 3D** scenes using graph-based alignment and cross-grained contrastive learning.

EXPERIENCES

Huawei Cooperation Project | *Project Leader*

Oct. 2023 – Jun. 2024

- Developed a **YOLO-MS**-based real-time algorithm for rapid and precise focus tracking.
- Deployed in the “**Feng Chi Shan Pai**” system of Huawei Mate 70 and P70.
- Recipient of the “**2024 Outstanding Technical Cooperation Project**” award from Huawei. [Report]

SenseTime | *Research Scientist Intern*

Oct. 2023 – Jun. 2024

- Real-LOD**: agentic workflow to improve LOD models from a data-centric perspective; accepted at **ICLR 2025**.
- Reproduced the OWLv2 training framework and implemented its training and evaluation on the REC task.

Huawei Cooperation Project | *Project Leader*

Jan. 2022 – Oct. 2022

- Proposed CNN-guided heterogeneous learning via **CrossKD**.
- Achieved SOTA detection performance with only 50% energy consumption using ANN-based inference.