Bin Ren

+39-329-324-8893 | bin.ren@unitn.it | https://amazingren.github.io/

Amazingren | Google Scholar | LinkedIn

Education

- **University of Pisa (UNIPI) & University of Trento (UNITN)** *11, 2021 - 10, 2025*
  National Ph.D. in AI for Society Italy
  - Supervised by: Prof. Dr. *Nicu Sebe* and Prof. Dr. *Rita Cucchiara*.
  - Thesis Title: Representation Learning Via Transformer: From 2D to 3D

- **Peking University (PKU)** *09, 2017 - 01. 2021*
  MS in Computer Applied Technology China
  - Supervised by: Prof. Dr. *Hong Liu*.
  - Thesis Title: Grouped Spatial-Temporal Enhancement Modeling Based Human Action Recognition
  - Awards: The 2nd prize of the May 4th Youth Science Award (Challenge Cup of PKU) in 06, 2019

- **Central South University (CSU)** *09, 2012 - 07, 2016*
  BA in Microelectronics Manufacturing Engineering China
  - Scholarship: Three times a 3rd-class scholarship winner in 2013, 2014, and 2015.

Visiting Research Experience

- **Visiting Ph.D. Student at INSAIT** *04,2024 - 06.2025*
  Advised by Prof. Dr. Luc Van Gool Sofia, Bulgaria
  - *Low-level Vision*:: At INSAIT, I led the development of efficient and scalable restoration models for diverse degradations (i.e., all-in-one). I proposed AnyIR (**submitted** to TMLR'25), a lightweight framework leveraging shared semantics for unified restoration with >80% reduction in parameters and FLOPs. Building on this, I developed MIRAGE (**submitted** to ICLR'26), which combines modular feature decomposition and manifold-based contrastive learning to enhance generalization. These works form a solid foundation for practical, high-quality image (and potentially video) enhancement under real-world conditions.
  - *Large Multimodal Models (LMM)*: I explored extending LMMs beyond natural images to structured, domain-specific modalities. We proposed EarthMind (**Submitted** to ICLR'26), a 4B vision-language framework for remote sensing, achieving SOTA performance on the proposed EarthMind-Bench and surpassing GPT-4o—demonstrating strong potential for generalization across multi-granular, multi-sensor visual domains.
  - *3D Representation Learning*: Additionally, I am interested in processing irregular 3D data such as point clouds and Gaussian splats, with a focus on extracting meaningful representations to support downstream tasks. I contributed to ShapeSplat (**Accepted** by 3DV'25 as oral), which introduces self-supervised learning on object-level 3D Gaussians via masked autoencoding. We then extended this to large-scale scenes with SceneSplat (**Accepted** by ICCV'25 as oral), addressing structured scene-level representation learning.

- **Visiting Ph.D. Student at Computer Vision Lab of ETH Zürich** *08, 2023 - 01, 2024*
  Advised by Prof. Dr. Luc Van Gool and in collaboration with Dr. Yawei Li Zürich, Switerland
  - *Low-Level Vision*: During my stay at the CVL of ETH Zürich, I proposed SemanIR, a transformer-based image restoration framework that constructs and shares key-semantic dictionaries to enable efficient and accurate restoration with linear attention complexity, achieving state-of-the-art performance across six benchmark tasks. This work has been **accepted** by NeurIPS'2024.
  - *3D Representation Learning*: Meanwhile, I also worked on a novel framework that integrates contrastive learning into masked autoencoding for 3D point cloud pretraining by using dual random masking and shared-weight reconstruction, achieving decent performance across classification, segmentation, and few-shot tasks. This work has been **accepted** by ACCV'2024.

PUBLICATIONS

C=CONFERENCE, J=JOURNAL, S=IN SUBMISSION, *=CORRESPONDING AUTHOR

[C.1] **Bin Ren**, Yawei Li, Jingyun Liang, Rakesh Ranjan, Mengyuan Liu, Rita Cucchiara, Luc Van Gool, Ming-Hsuan Yang, and Nicu Sebe. **Sharing Key Semantics in Transformer Makes Efficient Image Restoration**. Accepted by **NeurIPS** in 2024.

[C.2] **Bin Ren**, Yahui Liu, Yue Song, Wei Bi, Rita Cucchiara, Nicu Sebe, and Wei Wang. **Masked Jigsaw Puzzle: A Versatile Position Embedding for Vision Transformers**. Accepted by **CVPR** in 2023.

[C.3] **Bin Ren**, Guofeng Mei, Danda Pani Paudel, Weijie Wang, Yawei Li, Mengyuan Liu, Rita Cucchiara, Luc Van Gool, Nicu Sebe **Bringing masked autoencoders explicit contrastive properties for point cloud self-supervised learning**. Accepted by **ACCV** in 2024.

[C.4] **Bin Ren**, Hao Tang, Yiming Wang, Xia Li, Wei Wang, and Nicu Sebe. **PI-Trans: Parallel-ConvMLP and Implicit-Transformation Based GAN for Cross-View Image Translation**. Accepted by **ICASSP** in 2023.

[C.5] **Bin Ren**, Hao Tang, and Nicu Sebe. **Cascaded Cross MLP-Mixer GANs for Cross-View Image Translation**. Accepted by **BMVC** in 2021 as **Oral**.

[J.1] **Bin Ren**, Hao Tang, Fanyang Meng, Runwei Ding, Philip Torr, and Nicu Sebe. **Cloth interactive transformer for virtual try-on**. Accepted by **ACM ToMM** in 2023.

[S.1] **Bin Ren**, Eduard Zamfir, Zongwei Wu, Yawei Li, Yidi Li, Danda Pani Paudel, Radu Timofte, Ming-Hsuan Yang, Luc Van Gool, and Nicu Sebe. **Any Image Restoration via Efficient Spatial-Frequency Degradation Adaptation**. Submitted to **TMLR** in 2025.

[S.2] **Bin Ren**, Yawei Li, Xu Zheng, Yuqian Fu, Danda Pani Paudel, Ming-Hsuan Yang, Luc Van Gool, and Nicu Sebe. **Manifold-aware Representation Learning for Degradation-agnostic Image Restoration**. Submitted to **ICLR2026**.

[C.6] Mengjiao Ma, Qi Ma, Yue Li, Jiahuan Cheng, Runyi Yang, **Bin Ren***, Nikola Popovic, Mingqiang Wei, Nicu Sebe, Luc Van Gool, Theo Gevers, Martin R Oswald, Danda Pani Paudel. **SceneSplat++: A Large Dataset and Comprehensive Benchmark for Language Gaussian Splatting**. Accepted by **NeurIPS** in 2025.

[C.7] Yue Li, Qi Ma, Runyi Yang, Huapeng Li, Mengjiao Ma, **Bin Ren***, Nikola Popovic, Nicu Sebe, Ender Konukoglu, Theo Gevers, Luc Van Gool, Martin R Oswald, Danda Pani Paudel. **Scenesplat: Gaussian splatting-based scene understanding with vision-language pretraining**. Accepted by **ICCV** in 2025 as **Oral**.

[C.8] Yuqian Fu, Runze Wang, **Bin Ren**, Guolei Sun, Biao Gong, Yanwei Fu, Danda Pani Paudel, Xuanjing Huang, Luc Van Gool **ObjectRelator: Enabling Cross-View Object Relation Understanding Across Ego-Centric and Exo-Centric Perspectives**. Accepted by **ICCV** in 2025 as **Highlight**.

[C.9] Qi Ma, Yue Li, **Bin Ren***, Nicu Sebe, Ender Konukoglu, Theo Gevers, Luc Van Gool, and Danda Pani Paudel. **ShapeSplat: A Large-scale Dataset of Gaussian Splats and Their Self-Supervised Pretraining**. Accepted by **3DV** in 2025 as **Oral**.

[C.10] Mengyi Zhao, Mengyuan Liu, **Bin Ren**, Shuling Dai, and Nicu Sebe. **Modiff: Action-conditioned 3d motion generation with denoising diffusion probabilistic models**. Accepted by **ICASSP** in 2024.

[C.11] Chang Liu, Mengyi Zhao, **Bin Ren**, Mengyuan Liu, and Nicu Sebe. **Spatio-Temporal Graph Diffusion for Text-Driven Human Motion Generation**. Accepted by **BMVC** in 2023 as **Oral**.

[J.2] Wenhao Li, Mengyuan Liu, Hong Liu, **Bin Ren**, Xia Li, Yingxuan You, and Nicu Sebe. **HYRE: Hybrid Regressor for 3D Human Pose and Shape Estimation**. Accepted by **TIP** in 2024.

[S.3] Yan Shu, **Bin Ren***, Zhitong Xiong, Danda Pani Paudel, Luc Van Gool, Begum Demir, Nicu Sebe, Paolo Rota. **EarthMind: Towards Multi-Granular and Multi-Sensor Earth Observation with Large Multimodal Models**. Submitted to **ICLR2026**.

ACADEMIC SERVICES

- Workshop/Challenge Organizer: SPI (1st) at SWC'24, NTIRE (9th, 10th) at CVPR'(24,25), AIM (6th) at ICCV'25
- Conference Reviewer: NeurIPS, ICML, ICLR, CVPR, ICCV, ECCV, ACCV, BMVC, 3DV, BMVC, and ICASSP, etc.
- Journal Reviewer: TPAMI, TIP, IJCV, RAL, TMM, and ToMM, etc.
- Area Chair: BMVC'25

SKILLS

- Frequent user of **Python** and familiar with Linux OS, Git, Huggingface, and ViM.
- Frequent user of **PyTorch** and familiar with Deep Learning in Computer Vision.

ADDITIONAL INFORMATION

Languages: English (C1, Frequent use), Italian (A2, Basic), Mandarin (Native).
Interests: Cooking, Hiking, and Singing.