

# MINGXUAN LIU

+39-347-676-3600   [mingxuan.liu@unitn.it](mailto:mingxuan.liu@unitn.it)   [Google Scholar](#)   [Personal Web](#)   [Github](#)   [LinkedIn](#)

## TL;DR

I am a PhD student based in Italy, focusing in improving machine open-world perception, understanding, fairness and interpretability of visual input through multimodal reasoning and knowledge discovery.

## RESEARCH INTEREST

Agentic AI, Vision-Language, Multimodal Reasoning, Open-vocabulary Object Detection, Knowledge Discovery, Semi-/Un-supervised Learning, Layout2Img Diffusion Models, Large Language Models, Knowledge Retrieval

## EDUCATION

<b>University of California, Los Angeles</b> <i>Visiting Graduate Researcher on learning 4D scene graph for urban scene simulator; Advisor: Bolei ZHOU</i>	<b>Feb. 2025 – Jul. 2025</b> <i>Los Angeles, U.S.</i>
<b>Università di Trento</b> <i>PhD student in Deep Learning and Computer Vision; Advisor: Elisa RICCI</i>	<b>Nov. 2022 – Mar. 2026</b> <i>Trento, Italy</i>
<b>KTH Royal Institute of Technology</b> <i>MSc in Intelligent Autonomous Systems; Grade: A</i>	<b>Sep. 2021 – Jul. 2022</b> <i>Stockholm, Sweden</i>
<b>Università di Trento</b> <i>MSc in Mechatronics Engineering; Grade: 110L, Summa cum laude</i>	<b>Sep. 2020 – Aug. 2021</b> <i>Trento, Italy</i>
<b>Beijing University of Civil Engineering and Architecture</b> <i>BSc in Energy and Power Engineering</i>	<b>Sep. 2013 – Jun. 2017</b> <i>Beijing, China</i>

## WORK EXPERIENCE

<b>NAVER LABS Europe</b> <i>Visiting Researcher (hybrid), Visual Representation Learning Team</i> <ul style="list-style-type: none"><li>• <b>Topic:</b> 1) Improving open-vocabulary and vocabulary-free object detection on handling novel classes; 2) Benchmarking Layout2Image diffusion models.</li><li>• <b>Advisors:</b> Gabriela CSURKA, Riccardo VOLPI, and Tyler L. HAYES</li></ul>	<b>Jun. 2023 – Dec. 2024</b> <i>Grenoble, France</i>
<b>Ericsson</b> <i>Master Thesis Intern, Department of Device Software Research</i> <ul style="list-style-type: none"><li>• <b>Title:</b> The V-SLAM Hurdler: A Faster V-SLAM System using Online Semantic Dynamic-and-Hardness-aware Approximation</li><li>• Approximate spatial computing for Simultaneous Localization and Mapping (SLAM) algorithm on AR/MR devices</li><li>• Investigate quantization methods for CNN-based Object Detection (YOLOv4) and Instance Segmentation algorithms (Mask R-CNN)</li><li>• Investigate the cloud-device cooperation mechanism of distributed Semantic SLAM based on the approximate computing techniques</li></ul>	<b>Jan. 2022 – Jun. 2022</b> <i>Lund, Sweden</i>
<b>Siemens Co., Ltd.</b> <i>Innovation Engineer, Department of Innovation, Building Technology Division</i> <ul style="list-style-type: none"><li>• Conducted innovative application research, market analysis and competitive product analysis in building automation industry with Internet-of-Things (IoT) technology</li><li>• Designed software and hardware solution for IoT-based application and demonstrated the solution in a zero to one product definition and development fashion</li><li>• Defined features and functions for the innovative products based on the new use cases</li><li>• Cooperated with the internal R&amp;D and production departments, third-party partners (OEM manufacturers, value-added partners, system integrators) to develop and deploy innovative products</li><li>• Managed and deployed pilot projects of the innovative IoT-based products</li></ul>	<b>Jul. 2017 – Jul. 2020</b> <i>Beijing, China</i>

## PATENT

<b>A Method for Using Semantic Hierarchy Trees to Increase the Robustness of OvOD Models</b> US Patent App. (status: filed; under processing) Mingxuan Liu, Tyler Hayes, Elisa Ricci, Gabriela Csurka, Riccardo Volpi	<b>Mar. 2024</b>
---	------------------

## PUBLICATIONS

---

### Organizing Unstructured Image Collections using Natural Language

*Preprint*

Preprinted on Oct. 7 2024.

Keywords: Knowledge Discovery · Knowledge Retrieval · Agentic AI · LLM · VLM · Bias Discovery

**Mingxuan Liu**, Zhun Zhong, Jun Li, Gianni Franchi, Subhankar Roy, Elisa Ricci

### Test-time Vocabulary Adaptation for Language-driven Object Detection

*Under review*

Under review and awaiting U.S. patent filing prior to preprinting.

Keywords: Open-vocabulary Object Detection · LLM · VLM · Keyword Extraction

**Mingxuan Liu**, Tyler Hayes, Massimiliano Mancini, Elisa Ricci, Riccardo Volpi, Gabriela Csurka

### SHiNe: Semantic Hierarchy Nexus for Open-vocabulary Object Detection

*Seattle, U.S.*

The IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**; **accepted as a Highlight paper, 2.8% acceptance rate**), 2024.

Keywords: Open-vocabulary Object Detection · Semantic Hierarchy Retrieval · LLM · VLM

**Mingxuan Liu**, Tyler Hayes, Elisa Ricci, Gabriela Csurka, Riccardo Volpi

### Democratizing Fine-grained Visual Recognition with Large Language Models

*Vienna, Austria*

International Conference on Learning Representations (**ICLR**), 2024.

Keywords: Knowledge Discovery · Agentic AI · LLM · VLM · Fine-grained Visual Recognition

**Mingxuan Liu**, Subhankar Roy, Wenjing Li, Zhun Zhong, Nicu Sebe, Elisa Ricci

### Large-scale Pre-trained Models are Surprisingly Strong in Incremental Novel Class Discovery

*Kolkata, India*

International Conference on Pattern Recognition (**ICPR**; **accepted as an Oral paper**), 2024

Keywords: Novel Class Discovery · Representation Learning · Continual Learning

**Mingxuan Liu**, Subhankar Roy, Zhun Zhong, Nicu Sebe, Elisa Ricci

### Class-incremental Novel Class Discovery

*Tel Aviv, Israel*

European Conference on Computer Vision (**ECCV**), 2022.

Keywords: Novel Class Discovery · Semi-supervised Learning · Continual Learning

Subhankar Roy\*, **Mingxuan Liu**\*, Zhun Zhong, Nicu Sebe, Elisa Ricci (\*= equal contribution)

### Siemens RWG Control Platform Advanced Course and Practice

*Beijing, China*

Book; China Electric Power Press (**CEPP**), 2022; ISBN: 9787519859947

Keywords: PLC Controller · Low-code Programming · IoT · SaaS · Smart Building

Jiaxin Han, Huixia Zhao, Kaixuan Zhang, **Mingxuan Liu**

## GRANTS

---

- **OpenAI Researcher Access Program**, awarded 5,000 USD in API credits, Principal Investigator, 2024

- **ISCRA**, awarded 30,000 USD (8,000 V100 GPU Hours), Code: HP10C58YK9, Principal Investigator, Italy, 2022

## TECHNICAL SKILLS

---

**Programming language:** Python, C/C++, Ada

**Framework:** ROS, PyTorch, OpenCV, Point Cloud Library (PCL), G2O, Eigen, Ceres, Sophus, LaTeX, ORB SLAM

**Software:** Unity (Virtual RGB-D camera unity implementation), Maple (Kinematic and Dynamic simulation)

**Skills I am learning:** TensorRT, CUDA

## PERSONAL SKILLS

---

**Willing-to-learn:** New data structure, New modality, New tools, New fields

**Language proficiency:** Chinese (native), English (fluent), Italian (Ciao and Ti amo only)

**Hobbies:** Bodybuilding, Hiking, Bouldering, Cooking, Playing piano

## COMMUNITY SERVICE

---

**ICLR:** Reviewer 2025

**ICML:** Reviewer 2025

**NeurIPS:** Reviewer 2024

**CVPR:** Reviewer 2024, 2025

**ECCV:** Reviewer 2024

**Journals:** IJCV, CVIU