SHEN Xi

Experience

Intellindust

2023–present **Co-Founder and Chief Scientist**, Shenzhen.
Reliable Edge AI for Real-Time Vision Applications

Tencent AI Lab

2021–2023 **Senior Researcher**, Shenzhen.
Multimodal Generative Intelligence

Education

2017–2021 **Ph.D.**, *École des Ponts ParisTech*, Paris.
Thesis: Deep Learning for Near-duplicated Patterns Discovery and Alignment in Artworks
Advisor: Mathieu Aubry
Thesis committee: Patrick Pérez, Alexei A. Efros, Shiry Ginosar, Laurent Heutte, Yann Gousseau
Impact: my paper ArtMiner [3] was covered by Nature, my thesis is served as the basis for the ERC DISCOVER project

2013–2017 **Engineering programme**, *École des Ponts ParisTech*, Paris.
Department of Computer Science and Applied Mathematics

2016–2017 **Master**, *École Normale Supérieure Paris-Saclay*, Paris.
First-class honors (Mention très bien)
Mathematics, Computer Vision and Machine Learning (Mathématiques, Vision, Apprentissage)

2009–2013 **Dual Bachelor**, *Université Claude Bernard Lyon & Wuhan University*, Lyon & Wuhan.
First-class honors (Mention très bien)
Mechanics & Physics

Languages

Chinese Native **French** Fluent (10 years in France) **English** Fluent

Skills

Python (Pytorch), C++, Matlab, Latex
See my resleased code on GitHub: http://github.com/XiSHEN0220

Academic Services: Reviewer

Conference ICML, CVPR, ECCV, NeurIPS, ICCV, ICLR, BMVC, WACV
Journal IJCV, TPAMI, PR

Award

ICML 2022	Outstanding reviewer	*ICML Program Committe*
ECCV 2024	1st Place, Open-set Recognition challenge	*OOD Generalization in CV Workshop*
CVPR 2024	1st Place, Low-Light Object Detection Challenge	*Physics-Based Vision Workshop*
CVPR 2024	3rd Place, Low-Light Instance Segmentation Challenge	*Physics-Based Vision Workshop*
CVPR 2025	2nd Place, Joint Denoising and Demosaicing Challenge	*NTIRE*
CVPR 2025	2nd Place, Ball Action Spotting Challenge	*SoccerNet*
CVPR 2025	Honorable Mention, Foundation Few-shot Object Detection Challenge	*FSOD*

Teaching

2021 - 2021 Digital Humanities Meet Artificial Intelligence *Université Paris Sciences & Lettres*

shenxiluc@gmail.com • *http://xishen0220.github.io*

2021 - 2021	Computer Vision for Mechanics of Materials	*Ecole des Ponts ParisTech*
2019 - 2019	Signal Processing and Machine Learning	*Ecole des Ponts ParisTech*

Mentoring

2024 - 2025	Haiyang Xie	*Wuhan University*
2023 - 2024	Yuting Li	*China Three Gorges University*
2022 - 2023	Jianrong Zhang	*Jilin University*
2022 - 2023	Yangsong Zhang	*Shanghai Jiaotong University*
2020 - 2021	Oumayma Bounou	*Ecole des Ponts ParisTech*
2021 - 2022	Yingyi Chen	*KU Leuven*
2022 - 2022	Yangtao Wang	*Grenoble Computer Science Laboratory*
2022 - 2022	Wen Guo	*Inria RobotLearn*

Selected Publications

RED indicates serving as the corresponding author

First Author

[1] **Shen, Xi**, Robin Champenois, Shiry Ginosar, Ilaria Pastrolin, Morgane Rousselot, Oumayma Bounou, Tom Monnier, Spyros Gidaris, François Bougard, Pierre-Guillaume Raverdy, Marie-Françoise Limon, Christine Bénévent, Marc Smith, Olivier Poncet, K Bender, Joyeux-Prunel Béatrice, Elizabeth Honig, Alexei A Efros, and Mathieu Aubry. Spatially-consistent feature matching and learning for art collections and watermark recognition. *International Journal of Computer Vision (IJCV)*, 2022. **Project page:** http://imagine.enpc.fr/~shenx/HisImgAnalysis/.

[2] **Shen, Xi**, François Darmon, Alexei A Efros, and Mathieu Aubry. Ransac-flow: generic two-stage image alignment. In *European Conference on Computer Vision (ECCV)*, 2020. **Project page:** http://imagine.enpc.fr/~shenx/RANSAC-Flow/.

[3] **Shen, Xi**, Alexei A Efros, and Mathieu Aubry. Discovering visual patterns in art collections with spatially-consistent feature learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019. **Project page:** http://imagine.enpc.fr/~shenx/ArtMiner/.

[4] **Shen, Xi**, Alexei A Efros, Armand Joulin, and Mathieu Aubry. Learning co-segmentation by segment swapping for retrieval and discovery. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPRW)*, 2022. **Project page:** http://imagine.enpc.fr/~shenx/SegSwap/.

[5] **Shen, Xi**, Yang Xiao, Shell Xu Hu, Othman Sbai, and Mathieu Aubry. Re-ranking for image retrieval and transductive few-shot classification. In *Neural Information Processing Systems (NeurIPS)*, 2021. **Project page:** http://imagine.enpc.fr/~shenx/SSR/.

Corresponding Author

[6] Yingyi Chen, Shell Xu Hu, **Shen, Xi**, Chunrong Ai, and Johan A. K. Suykens. Compressing features for learning with noisy labels. *IEEE Transactions on Neural Networks and Learning Systems (TNNLS)*, 2022. **Project page:** https://yingyichen-cyy.github.io/CompressFeatNoisyLabels/.

[7] Yingyi Chen, **Shen, Xi**, Yahui Liu, Qinghua Tao, and Johan A. K. Suykens. Jigsaw-vit: Learning jigsaw puzzles in vision transformer. *Pattern Recognition Letters*, 2022. **Project page:** https://yingyichen-cyy.github.io/Jigsaw-ViT/.

[8] Wen Guo, Yuming Du, **Shen, Xi**, Vincent Lepetit, Alameda-Pineda Xavier, and Moreno-Noguer Francesc. Back to mlp: A simple baseline for human motion prediction. In *Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2022. **Code:** https://github.com/dulucas/siMLPe.

[9] Shihua Huang, Yongjie Hou, Longfei Liu, Xuanlong Yu, and **Shen, Xi**. Real-time object detection meets dinov3. *arXiv*, 2025. **Project page:** https://intellindust-ai-lab.github.io/projects/DEIMv2/.

[10] Shihua Huang, Zhihao Lu, Xiaodong Cun, Yongjun Yu, Xiao Zhou, and **Shen, Xi**. Deim: Detr with improved matching for fast convergence. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025. **Project page:** https://www.shihuahuang.cn/DEIM/.

[11] Yuting Li, Dexiong Chen, Tinglong Tang, and **Shen, Xi**. Htr-vt: Handwritten text recognition with vision transformer. *Pattern Recognition*, 2024. **Project page:** https://yutingli0606.github.io/HTR-VT/.

[12] Yuting Li, Yingyi Chen, Xuanlong Yu, Dexiong Chen, and **Shen, Xi**. Sure: Survey recipes for building reliable and robust deep networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024. **Project page:** https://yutingli0606.github.io/SURE/.

[13] Yangtao Wang, **Shen, Xi**, Shell Xu Hu, Yuan Yuan, James Crowley, and Dominique Vaufreyday. Self-supervised transformers for unsupervised object discovery using normalized cut. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. **Project page:** https://www.m-psi.fr/Papers/TokenCut2022/.

[14] Yangtao Wang, **Shen, Xi**, Yuan Yuan, Yuming Du, Maomao Li, Shell Xu Hu, James L Crowley, and Dominique Vaufreydaz. Tokencut: Segmenting objects in images and videos with self-supervised transformer and normalized cut. In *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2023. **Project page:** https://www.m-psi.fr/Papers/TokenCut2022/.

[15] Jianrong Zhang, Yangsong Zhang, Xiaodong Cun, Shaoli Huang, Yong Zhang, Hongwei Zhao, Hongtao Lu, and **Shen, Xi**. T2M-GPT: Generating Human Motion from Textual Descriptions with Discrete Representations. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. **Project page:** https://mael-zys.github.io/T2M-GPT/.

Collaboration

[16] Xie Haiyang, **Shen, Xi**, Huang Shihua, Wang Qirui, and Wang Zheng. Simrod: A simple baseline for raw object detection with global and local enhancements. In *The Association for the Advancement of Artificial Intelligence (AAAI)*, 2026. **Project page:** https://ocean146.github.io/SimROD2025/.

[17] Shell Xu Hu, Pablo G Moreno, Yang Xiao, **Shen, Xi**, Guillaume Obozinski, Neil D Lawrence, and Andreas Damianou. Empirical bayes transductive meta-learning with synthetic gradients. In *International Conference on Learning Representations (ICLR)*, 2019. **Code:** https://github.com/hushell/sib_meta_learn.

[18] Lang Rao, Yuan Yuan, **Shen, Xi**, Yu Guocan, and Chen Xiaoyuan. Designing nanotheranostics with machine learning. *Nature Nanotechnology*, 2024.

[19] Liu Weihuang, **Shen, Xi**, Pun Chi-Man, and Cun Xiaodong. Explicit visual prompting for low-level structure segmentations. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. **Project page:** https://nifangbaage.github.io/Explict-Visual-Prompt/.

[20] Liu Weihuang, **Shen, Xi**, Pun Chi-Man, and Cun Xiaodong. Explicit visual prompting for universal foreground segmentations. In *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2023. **Project page:** https://nifangbaage.github.io/Explict-Visual-Prompt/.

[21] Liu Weihuang, **Shen, Xi**, Li Haolun, Bi Xiuli, Liu Bo, Pun Chi-Man, and Cun Xiaodong. Depth-aware test-time training for zero-shot video object segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024. **Project page:** https://nifangbaage.github.io/DATTT/.

[22] Zhang Wenxuan, Cun Xiaodong, Wang Xuan, Zhang Yong, **Shen, Xi**, Guo Yu, Shan Ying, and Wang Fei. Sadtalker: Learning realistic 3d motion coefficients for stylized audio-driven single image talking

face animation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. **Project page:** https://sadtalker.github.io/.

[23] Zhi Yihao, Cun Xiaodong, Chen Xuelin, **Shen, Xi**, Guo Wen, Huang Shaoli, and Gao Shenghua. Livelyspeaker: Towards semantic-aware co-speech gesture generation. In *International Conference on Computer Vision (ICCV)*, 2024. **Project page:** https://nifangbaage.github.io/Explict-Visual-Prompt/.

[24] Yuan Yuan, Yueming Lyu, **Shen, Xi**, Ivor W Tsang, and Dit-Yan Yeung. Marginalized average attentional network for weakly-supervised learning. In *International Conference on Learning Representations (ICLR)*, 2019. **Code:** https://github.com/yyuanad/MAAN.