

# Satoshi Tsutsui

Postdoctoral Research Fellow at ROSE Lab, Nanyang Technological University, Singapore.  
satoshi.tsutsui [at] ntu.edu.sg  
<https://hellosatoshi.github.io>  
Last Update: Mar 2024

## Education

- Ph.D., School of Informatics, Computing, and Engineering, Indiana University, Bloomington, USA. 2021
  - Dissertation: *Rethinking the Role of Training Data for Computer Vision: Scientific Studies of Egocentric Vision*.
  - Advisors: David Crandall (committee chair) and Chen Yu (committee member).
- M.S., School of Informatics, Computing, and Engineering, Indiana University, Bloomington, USA. 2017
  - Advisor: Ying Ding.
- B.E. (with highest honors), Faculty of Science and Technology, Keio University, Tokyo, Japan. 2015

## Experience

- Postdoctoral Research Fellow at Nanyang Technological University, Singapore. September 2022 - Present
  - Advisor: Bihan Wen. Lab Director: Alex Kot.
- Postdoctoral Research Fellow at National University of Singapore. December 2021 - September 2022
  - Advisor: Mike Shou.
- Research Intern at Facebook, USA. December 2020 - June 2021
  - Mentor: Ruta Desai and Karl Ridgeway.
  - Developed visual perception algorithms for AR/VR devices.
- Visiting Ph.D. Student at Fudan University, China. May 2019 - August 2019
  - Mentor: Yanwei Fu.
  - Worked on few-shot visual recognition.
- Visiting Ph.D. Student at Peking University, China. May 2018 - August 2018
  - Mentor: Liangcai Gao.
  - Worked on computer vision for medical images.
- Research Intern at Preferred Networks, Japan. May 2017 - August 2017
  - Mentor: Tommi Kerola and Shunta Saito.
  - Developed semantic segmentation algorithms for autonomous driving.

## Publications

1. Xueyi Ke, Satoshi Tsutsui, Yayun Zhang, and Bihan Wen. (2025). Discovering Hidden Visual Concepts Beyond Linguistic Input in Infant Learning. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. (Acceptance rate =  $2878/13008 = 22.12\%$ ).
2. Xiyu Wang, Yufei Wang, Satoshi Tsutsui, Weisi Lin, Bihan Wen, and Alex Kot. (2024). Evolving Storytelling: Benchmarks and Methods for New Character Customization with Diffusion Models. *ACM International Conference on Multimedia (ACMMM)*. (Oral acceptance rate =  $174/4385 = 3.97\%$ , Overall acceptance rate =  $1150/4385 = 26.23\%$ ).
3. Winnie Pang, Xueyi Ke, Satoshi Tsutsui, and Bihan Wen. (2024). Integrating Clinical Knowledge into Concept Bottleneck Models. *Medical Image Computing and Computer Assisted Intervention (MICCAI)*. (Acceptance rate =  $858/2771 = 30.96\%$ ).
4. Juan Hu, Xin Liao, Difei Gao, Satoshi Tsutsui, Qian Wang, Zheng Qin, and Mike Zheng Shou. (2024). Delocate: Detection and Localization for Deepfake Videos with Randomly-Located Tampered Traces. *International Joint Conference on Artificial Intelligence (IJCAI)*.

5. Satoshi Tsutsui, Winnie Pang, and Bihan Wen. (2023). WBCAtt: A White Blood Cell Dataset Annotated with Detailed Morphological Attributes. *Advances in Neural Information Processing Systems (NeurIPS)*. (Acceptance rate =  $322/987 = 32.62\%$ ).
6. Xizhe Xue, Dongdong Yu, Lingqiao Liu, Yu Liu, Satoshi Tsutsui, Ying Li, Zehuan Yuan, Ping Song, and Mike Zheng Shou. (2023). Transformer-based Open-world Instance Segmentation with Cross-task Consistency Regularization. *ACM International Conference on Multimedia (ACMMM)*. (Acceptance rate =  $902/3072 = 29.3\%$ ).
7. Alex Jinpeng Wang, Yixiao Ge, Rui Yan, Yuying Ge, Kevin Qinghong Lin, Satoshi Tsutsui, Xudong Lin, Guanyu Cai, Jianping Wu, Ying Shan, Xiaohu Qie, and Mike Zheng Shou. (2023). All in One: Exploring Unified Video-Language Pre-training. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (Acceptance rate =  $2360/9155 = 25.8\%$ ).
8. Satoshi Tsutsui, Zhengyang Su, and Bihan Wen. (2023). Benchmarking White Blood Cell Classification Under Domain Shift. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*.
9. Zan-Xia Jin, Mike Zheng Shou, Fang Zhou, Satoshi Tsutsui, Jingyan Qin, and Xu-Cheng Yin. (2022). From Token to Word: OCR Token Evolution via Contrastive Learning and Semantic Matching for Text-VQA. *ACM International Conference on Multimedia (ACMMM)*. (Acceptance rate =  $690/2473 = 27.9\%$ ).
10. Eric Zhongcong Xu, Zeyang Song, Satoshi Tsutsui, Chao Feng, Mang Ye, and Mike Zheng Shou. (2022). AVA-AVD: Audio-visual Speaker Diarization in the Wild. *ACM International Conference on Multimedia (ACMMM)*. (Acceptance rate =  $690/2473 = 27.9\%$ ).
11. Satoshi Tsutsui, Xizi Wang, Guangyuan Weng, Yayun Zhang, Chen Yu, and David Crandall. (2022). Action Recognition based on Cross-Situational Action-object Statistics. *International Conference on Development and Learning (ICDL)*.
12. Satoshi Tsutsui, Yanwei Fu, and David Crandall. (2022). Reinforcing Generated Images via Meta-learning for One-Shot Fine-Grained Visual Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*. (Impact Factor = 24.31).
13. Satoshi Tsutsui, Yanwei Fu, and David Crandall. (2021). Whose hand is this? Person Identification from Egocentric Hand Gestures. *Winter Conference on Applications of Computer Vision (WACV)*. (First round acceptance; Acceptance rate =  $496/1241 = 35.4\%$ ).
14. Satoshi Tsutsui, David Crandall, and Chen Yu. (2021). Reverse-engineer the Distributional Structure of Infant Egocentric Views for Training Generalizable Image Classifiers. *International Workshop on Egocentric Perception, Interaction and Computing (EPIC)*, In conjunction with the *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. (Extended Abstract).
15. Satoshi Tsutsui, Ruta Desai, and Karl Ridgeway. (2021). How You Move Your Head Tells What You Do: Self-supervised Video Representation Learning with Egocentric Cameras and IMU Sensors. *International Workshop on Egocentric Perception, Interaction and Computing (EPIC)*, In conjunction with the *IEEE International Conference on Computer Vision (ICCV)*. (Extended Abstract).
16. Satoshi Tsutsui, Arjun Chandrasekaran, Md Reza, David Crandall, and Chen Yu. (2020). A Computational Model of Early Word Learning from the Infant’s Point of View. *Annual Conference of the Cognitive Science Society (CogSci)*. (Oral Acceptance Rate =  $177/811 = 22\%$ ).
17. Satoshi Tsutsui, Yanwei Fu, and David Crandall. (2019). Meta-Reinforced Synthetic Data for One-Shot Fine-Grained Visual Recognition. *Advances in Neural Information Processing Systems (NeurIPS)*. (Poster Acceptance Rate =  $1428/6743 = 21\%$ ).
18. Satoshi Tsutsui, Dian Zhi, Md Alimoor Reza, David Crandall, and Chen Yu. (2019). Active Object Manipulation Facilitates Visual Object Learning: An Egocentric Vision Study. *International Workshop on Egocentric Perception, Interaction and Computing (EPIC)*, In conjunction with the *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. (Extended Abstract).
19. Zheng Gao, Gang Fu, Chunping Ouyang, Satoshi Tsutsui, Xiaozhong Liu, Jeremy Yang, Christopher Gessner, Brian Foote, David Wild, and Ying Ding. (2019). edge2vec: Representation Learning Using Edge Semantics for Biomedical Knowledge Discovery. *BMC Bioinformatics*. 20(1), 306. (Impact Factor = 3.213).

20. Satoshi Tsutsui, Tommi Kerola, Shunta Saito, and David J Crandall. (The first three authors have equal contribution). (2018). Minimizing Supervision for Free-space Segmentation. *Workshop on Autonomous Driving (WAD), In conjunction with the Conference on Computer Vision and Pattern Recognition (CVPR)*.
21. Satoshi Tsutsui, Sven Bambach, David Crandall, and Chen Yu. (2018). Estimating head motion from egocentric vision. *ACM International Conference on Multimodal Interaction (ICMI)*.
22. Ting-Ting Liang, Mengyan Sun, Liangcai Gao, Jing-Jing Lu, and Satoshi Tsutsui. (2018). APNet: semantic segmentation for pelvic MR image. *Chinese Conference on Pattern Recognition and Computer Vision (PRCV)*.
23. Satoshi Tsutsui, Zheng Gao, Yuzhuo Wang, Guilin Meng, and Ying Ding. (2018). A case study on viziometrics: What's the role of western blots in Alzheimer's Disease literature?. *iConference*. (Poster).
24. Satoshi Tsutsui, and David J Crandall. (2017). A data driven approach for compound figure separation using convolutional neural networks. *IAPR International Conference on Document Analysis and Recognition (ICDAR)*. (Oral Acceptance Rate =  $52/409 = 13\%$ ).
25. Satoshi Tsutsui, Tommi Kerola, and Shunta Saito. (2017). Distantly supervised road segmentation. *Workshop on Computer Vision for Road Scene Understanding and Autonomous Driving (CVRSUAD), In conjunction with the IEEE International Conference on Computer Vision (ICCV)*.
26. Baitong Chen, Satoshi Tsutsui, Ying Ding, and Feicheng Ma. (2017). Understanding the topic evolution in a scientific domain: An exploratory study for the field of information retrieval. *Journal of Informetrics*. 11(4), 1175-1189. (Impact Factor = 3.879).
27. Satoshi Tsutsui, Guilin Meng, Xiaohui Yao, David Crandall, and Ying Ding. (2017). Analyzing Figures of Brain Images from Alzheimer's Disease Papers. *iConference*. (Poster).
28. Satoshi Tsutsui, and David Crandall. (2017). Using artificial tokens to control languages for multilingual image caption generation. *Language and Vision Workshop, In conjunction with the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. (Extended Abstract).
29. Satoshi Tsutsui, Yi Bu, and Ying Ding. (2017). Using machine reading to understand Alzheimer's and related diseases from the literature. *Journal of Data and Information Science*. 2(4), 81-94. (Impact Factor = 1.771).
30. Satoshi Tsutsui, Ying Ding, and Guilin Meng. (2016). Machine reading approach to understand Alzheimer's disease literature. *International Workshop on Data and Text Mining in Biomedical Informatics (DTMBIO), In conjunction with the ACM Conference on Information and Knowledge Management (CIKM)*.

## Awards

- |   |             |
|---|-------------|
| • NeurIPS student travel award.                                     | 2019        |
| • ICDAR student travel award.                                       | 2017        |
| • Scholarship for Study Abroad from Yoshida Scholarship Foundation. | 2015 - 2017 |
| • Keio University Scholarship for Excellent Undergraduate Students. | 2012        |