

Vijay John

CONTACT INFORMATION	<i>Office Address:</i> Lawrence Technological Address, <i>Mobile:</i> +1 248-996-5341 21000 West Ten Mile Road Southfield, MI, 48075- 1058 <i>E-mail:</i> vjohn@ltu.edu	
RESEARCH FIELD	Multimodal Learning, Sensor Fusion (Data Fusion, Deep Learning-based Fusion, Bayesian Inference), Robotics (Autonomous Systems, Robot Perception, Human-Robot Interaction), Intelligent Mobility (Autonomous Vehicles, Wheeled Robots), Computer Vision and Machine Learning (Supervised, Unsupervised, Representation Learning, Weak Supervised Learning, Transfer Learning), Metric Learning (Distance Metric Optimization, Embedding Learning), Probabilistic Machine Learning	
RESEARCH INTEREST	Deep Learning (Neural Networks, CNNs, RNNs, Transformer Models), Missing Modality-based Multimodal Learning (Modality Imputation, Cross-Modal Learning), Perception in Humanoid Robotics (Visual Perception, Multimodal Perception, Robotic Vision), Gesture and Sign Language Recognition (Action Recognition, Pose Estimation, Temporal Sequence Analysis, Spatio-Temporal Learning)	
EDUCATION	University of Dundee , Dundee, Scotland, UK <i>School of Computing</i> Ph.D. in Computer Science, 2011 • Dissertation: “Markerless Multiple-View Human Motion Analysis Using Swarm Optimization and Subspace Learning” • Advisor: Prof. Emanuele Trucco Carnegie Mellon University , Pittsburgh, Pennsylvania, USA <i>Robotics Institute</i> M.S. in Information Technology and Robotics, 2007 • Dissertation: “Fingerprint-Based Biometric Authentication System” • Advisor: Prof. Mel Seigel Anna University , Chennai, India <i>Sri Sivasubramaniya Nadar College of Engineering</i> B.E. in Electronics and Communication Engineering, 2005 • Dissertation: “Verilog Design and Synthesis of a 64-Bit Microprocessor”	
ACADEMIC APPOINTMENTS	Lawrence Technological University , Southfield, USA <i>Tenure-track Associate Professor of Computer Science</i>	August 2025 – Current
	RIKEN , Kyoto, Japan <i>Research Scientist</i>	September 2021 – August 2025
	Toyota Technological Institute , Nagoya, Japan <i>Assistant Professor</i>	July 2017 – March 2021
	Toyota Technological Institute , Nagoya, Japan <i>Commissioned Scientist</i>	February 2017 – June 2017
	Toyota Technological Institute , Nagoya, Japan <i>Postdoctoral Researcher</i>	February 2014 – February 2017

Toyota Technological Institute, Chicago, USA
Visiting Researcher

March 2015

University of Amsterdam, Amsterdam, The Netherlands
Postdoctoral Researcher

2011 – 2013

HONORS AND AWARDS

Awards

Most Influential Paper of the Decade, 19th International Conference on Machine Vision Applications, July 2025

V. John, Z. Liu, S. Mita, B. Qi “Pedestrian Detection in Thermal Images Using Adaptive Fuzzy C-Means Clustering and Convolutional Neural Networks”, International Conference on Machine Vision Applications, 2015, pp. 246-249, doi: 10.1109/MVA.2015.7153177

Toyoda Award for the Best Researcher, Toyota Technological Institute

V. John. “Deep Learning for Intelligent Vehicle Applications.” 2016.

Best Paper Awards

T. Nguyen, Y. Kawanishi, V. John, T. Komamizu, I. Ide, “MultiSensor-Home: A Wide-area Multimodal Multi-view Dataset for Action Recognition and Transformer-based Sensor Fusion”, IEEE International Conference on Automatic Face and Gesture Recognition, 2025

V. John, Z. Liu, S. Mita. “Robust Traffic Light and Arrow Detection using Optimal Camera Parameters and GPS-based Prior.” IEEE Asia-Pacific Conference on Intelligent Robot Systems, 2016.

V. John, Z. Liu, C. Guo, S. Mita, K. Kidono. “Real-Time Lane Estimation Using Deep Features and Extra Trees Regression.” Pacific Rim Symposium on Image and Video Technology, 2015.

V. John, E. Trucco, S. McKenna. “Marker-less Human Motion Capture using Charting and Manifold-Constrained Particle Swarm Optimization.” British Machine Vision Conference Workshop, 2010.

Press

Ph.D. research featured in *Science Scotland*, Royal Society of Edinburgh.

RESEARCH PROJECTS

GUARDIAN ROBOT RESEARCH. 2021 - 2025

The project is carried out at RIKEN. Six teams are involved in developing a social robot with human-like characteristics for various applications.

Missing Modality Problem in Multimodal Learning

Research Problem: Deep learning-based multimodal perception report state-of-the-art accuracy. However, existing frameworks require the presence of all modalities, and this approach is limited by the problem of missing modalities, where one or more of the modalities could be missing. This is the missing modality problem. In this research, we address the missing modality problem.

Key Contributions:

- Addressed the missing modality problem in multimodal person classification by proposing a multimodal cascaded framework with three deep learning models and metric learning.

- Developed a method to generate complete multimodal data from incomplete inputs in the feature space via a latent space.
- Identified and addressed the multimodal portability problem, enabling the model to classify using any subset of modalities. Proposed KModNet to tackle both the missing modality and multimodal portability problems.
- Proposed a new audio-visual person recognition framework that enhances robustness in the absence of visual data. Utilized audio-based person attributes and a multi-head attention transformer to improve recognition accuracy under missing modality conditions.

Multimodal Weak Supervised Learning

Research Problem: This research on weak supervised multimodal learning addresses the challenge of training a video-based action recognition model using weak labels, specifically sequence-level annotations instead of complex frame-level labels.

Key Contributions:

- Proposed a novel learning framework for weakly supervised multimodal learning.
- Base model initially trained with weak labels to generate latent embeddings using an unique latent loss function.
- Resulting view-specific embeddings enhance downstream models for action recognition and detection tasks.

Multimodal Perception

Research Problem: This research focuses on multimodal perception for human-robot interactions using visible cameras, thermal cameras, and microphones. We aim to enhance the classification accuracy of visible-thermal person classification and improve emotion recognition through advanced deep learning techniques.

Key Contributions:

- Proposed a visible-thermal person classifier utilizing transfer learning, knowledge distillation, and vision transformers.
- Implemented visible and thermal teacher models to guide the multimodal classifier using a novel loss function for knowledge distillation.
- Developed a transformer-based model for audio-visual emotion recognition, featuring three transformer branches: audio self-attention, video self-attention, and audio-video cross-attention.

MULTIMODAL LEARNING FOR INTELLIGENT MOBILITY APPLICATIONS. 2014-2021.

These projects were carried out as a collaborative research between Toyota Technological Institute, Denso, Japan, Nippon Soken, Japan, Aisin, Japan and Toyota R&D Labs, Japan and Tier IV. I was the principal contributor in all the aforementioned projects.

Multimodal Perception:

Research Problem: This research enhanced multimodal perception for autonomous vehicles through the integration of diverse sensor modalities, including visible cameras, thermal cameras, milliwave radar, stereo cameras, and LIDAR

Key Contributions:

- Developed novel deep sensor fusion frameworks (RVNet and SO-Net) for integrating radar and appearance information to enhance vehicle detection and semantic segmentation.
- Formulated deep sensor fusion methods for visible and thermal cameras to improve semantic segmentation, pedestrian behavior estimation, scene forecasting, and semantic optimization.

- Single shot object detection, multi-label multi-class semantic segmentation frameworks, adversarial learning and latent variable modeling used to predict vehicle and pedestrian intention in driving scenes.
- Investigated the integration of LIDAR and stereo camera depth information without calibration aids to enhance depth precision.
- Created the ChiNet and PsiNet framework for fusing depth and appearance information from stereo camera pairs, advancing multimodal perception capabilities to predict the lane markers even under conditions of occlusion, shadows and missing lane markers.

Gesture Recognition and Vehicle Localization

Research Problem: This research advances touchless automotive user interfaces by integrating deep learning and sparse modeling techniques for gesture recognition from video data. Additionally, the project will focus on the localization of autonomous vehicles using 3D point cloud data obtained from stereo vision-based depth information.

Key Contributions:

- Developed a deep learning and sparse modeling framework for recognizing gestures from video, enabling touchless interaction with automotive user interfaces.
- Designed a robust system for intuitive user gesture recognition that enhances the overall user experience in vehicles.
- Introduced a novel particle swarm optimization framework combined with Kalman filtering for the localization of autonomous vehicles using stereo vision-based depth information.
- Generated virtual depth images from 3D point clouds to improve the accuracy and reliability of vehicle localization in complex environments.

AUTOMATIC CALIBRATION OF MULTIPLE NON-OVERLAPPING FOV CAMERAS. 2011-2013.

The project was carried out as a collaborative research between University of Amsterdam, Philips Research and Eagle Vision, Netherlands.

Research Problem: The research problem addressed the extrinsic calibration of camera networks with non-overlapping fields of view, addressing challenges such as varying illumination and camera gain.

Key Contributions:

- Developed a novel probabilistic approach for appearance-based person reidentification in non-overlapping camera networks.
- Introduced a graphical model that accounts for varying illumination and camera gain, alongside person appearance.
- Derived analytical solutions for person appearance and camera properties within the model.
- Proposed an automatic camera calibration algorithm using multiple trajectories and particle swarm optimization for multi-camera networks.
- Refined calibration parameters using a Bayesian framework with the Metropolis algorithm to correct accumulation errors.

HUMAN MOTION ANALYSIS USING MULTIPLE CAMERAS. 2007-2011.

PhD Thesis: My PhD thesis was carried out at the School of Computing of the University of Dundee concerning the markerless human motion tracking and classification of multiple-view video sequences using particle swarm optimisation and charting.

Key Contributions:

- Developed a method for markerless human motion tracking using multiple-view video sequences.
- Utilized particle swarm optimization and charting for efficient human motion analysis.
- Proposed a framework for human motion classification, identifying action types from extracted motion data.
- Targeted practical applications of human motion analysis across multiple domains.
- I was also responsible for the maintenance and update of the motion capture studio, which contained in-house developed acquisition software, multiple synchronized camera and professional lighting control.

Course Coordinator

Coordinating a machine learning and computer vision-based autonomous driving course for the Nodal Robotics Center in Government Engineering College, Thrissur, India in 2018.

Lecturer

- Datastructures, Lawrence Technological University, Fall 2025 - Current
- Computer Vision and Scene Understanding, Lawrence Technological University, Fall 2025 - Current
- Multimodal Learning, Lawrence Technological University, Spring 2026
- Collaborative Research Projects I and II, Fall 2025 - Current
- Introduction to Robotics and Intelligent Mobility, REC University, Spring 2025
- Understanding Deep Learning, SSN University, 2024
- Introduction to Robotics and Intelligent Mobility, SSN University, 2023
- Advances in Intelligent Mobility, SSN University, 2022
- Introduction to Intelligent Mobility, SSN University, 2021-2022

Co-lecturer

- English Writing, Toyota Technological Institute, 2018- 2021
Course coordinator: Prof. M. Okamoto
- English for Academic Presentation, Toyota Technological Institute, 2017-2021
Course coordinator: Nicholas Yaxley
- Text Mining and Collective Intelligence, University of Amsterdam, 2012
Course coordinator: Dr. Hayley Hung

Teaching assistant

- Research Methods, University of Dundee, 2011
Course coordinator: Dr. Zhang
- Signals and Images, University of Dundee, 2007-2011
Course coordinators: Dr. Jesse Hoey and Dr. Keith Edwards
- Software Development using C++, University of Dundee, 2007-2008
Course coordinator: Dr. Norman Alm

Course work

- Worked with Prof. Mel Siegel to review and repair the course materials for the “Sensing and Sensors” course in Robotics Institute in Carnegie Mellon University, 2006-2007.

SUPERVISION

Current Students

- PhD student **Nguyen Trung Thanh**, who is working on multimodal learning using multi-modal multi-view perception.

Past Students

- Research assistant, **Annamalai Lakshmanan**, who was working on deep learning-based semantic forecasting.
- Post-doctoral Research assistant, **Swarn Singh Rathour**, who was working on deep learning-based autonomous driving. Currently working as Researcher in Hitachi, Japan.
- PhD Student, **Weifu Li**, who was working on deep learning-based medical imaging.
- Master students: **Ryan Tocher** (animating human pose using Poser), **Leon Hashimoto** (person re-identification using RGB-D camera), **Jabez David** (monitoring the growth of bio-fuel algal cells using a microcontroller), **Mariko Umetsu** (deep learning-based gesture recognition) **Prareek Patel**, **Fakheerudin Mohammad**, and **Sunandana**
- Bachelors student, **Shogo Tsuchizawa** who was working on thermal and visible camera registration and pedestrian detection. Currently working as engineer in Daikin, Japan **Selicia S.**, who was working on deep learning-based tamil sign language recognition. **Jiayi Chen**, who was working on multimodal emotion recognition. **Yo Hsin Fang**, who was working on multimodal gesture generation.
- Research assistant, **Lyu Zheming**, who was working on the traffic light detection. Currently a Masters student in Toyota Technological Institute, Japan
- Research assistant, **Meenakshi Nithlan**, who was working on deep learning-based autonomous driving.

FUNDING

Industrial funding from Tier-IV

Pedestrian Behavior Prediction using Deep Generative Modelling and Deep Learning
2019-2020, US\$30,000
Role: Principle investigator.

Industrial funding from Denso

Investigation of Deep learning for Perception, Path Planning and Vision-based Controls, 2015-2019, US\$250,000
Role: Co-principle investigator.

Industrial funding from Aisin Seki

Investigation of Deep learning for Automated Driving, 2018-2019, US\$50,000
Role: Co-principle investigator.

Industrial funding from Toyota Central, R&D Labs

Investigation of Deep learning for Road Scene Understanding and Lane Estimation, 2015-2018, US\$60,000
Role: Co-principle investigator.

Industrial funding from Nippon Soken

Investigation of Deep learning for On-board Gesture Recognition, 2016-2017, US\$50,000
Role: Co-principle investigator.

Toyota Technological Institute Travel Grant

Visiting Toyota Technological Institute, Chicago, USA for collaborative research with Dr. Greg Shakhnarovich.

PATENTS

1. **Vehicle Control Device.** Inventors: Misa Harada, Yusuke Tanaka, Teraji, Seiichi Mita, Vijay John. Applicant/Owner: SOKEN Co., Ltd., DENSO Corporation, Toyota School Foundation. *Patent No. JP7303521B2*, Japan. Filed: June 28, 2019. Published: January 28, 2021. Application No.: 2019-121431. Publication No.: 2021-009440.
2. **Steering Angle Determination Device and Autonomous Vehicle.** Inventors: Kazutoshi Ishimaru, Tehrani Nikinejad Hossein, Masataka Konishi, Vijay John, Seiichi Mita. Applicant/Owner: SOKEN Co., Ltd., DENSO Corporation, Toyota School Foundation. *Patent No. JP7078458B2*, Japan. Filed: May 30, 2018. Published: December 5, 2019. Application No.: 2018-103663. Publication No.: 2019-206310.
3. **Recognition Device.** Inventors: Chen Bo-hsiang, Tehrani Nikinejad Hossein, Vijay John, Seiichi Mita, Sakiko Nishino, Kazutoshi Ishimaru. Applicant/Owner: DENSO Corporation, Toyota School Foundation, SOKEN Co., Ltd. *Patent No. JP7071215B2*, Japan. Filed: May 24, 2018. Published: November 28, 2019. Application No.: 2018-099724. Publication No.: 2019-204338.
4. **Steering Angle Determination Device, Autonomous Vehicle.** Inventors: Kazutoshi Ishimaru, Tehrani Nikinejad Hossein, Vijay John, Seiichi Mita. Applicant/Owner: SOKEN Co., Ltd., DENSO Corporation, Toyota School Foundation. *Patent No. JP6923362B2*, Japan. Filed: May 30, 2017. Published: December 20, 2018. Application No.: 2017-106180. Publication No.: 2018-199473.
5. **Position Estimation Device, Mobile Device.** Inventors: Sakiko Nishino, Tehrani Nikinejad Hossein, Kazutoshi Ishimaru, Shu Yu-Chuan, Vijay John, Seiichi Mita. Applicant/Owner: SOKEN Co., Ltd., DENSO Corporation, Toyota School Foundation. *Patent No. JP6959032B2*, Japan. Filed: May 17, 2017. Published: December 6, 2018. Application No.: 2017-098006. Publication No.: 2018-194417.
6. **Boundary Line Estimation Device.** Inventors: Kazutoshi Ishimaru, Tehrani Nikinejad Hossein, Takamasa Shirai, Vijay John, Seiichi Mita. Applicant/Owner: SOKEN Co., Ltd., DENSO Corporation, Toyota School Foundation. *Patent No. JP6798860B2*, Japan. Filed: November 29, 2016. Published: June 7, 2018. Application No.: 2016-231472. Publication No.: 2018-088151.
7. **Recognition Device, Program.** Inventors: Kazutoshi Ishimaru, Tehrani Nikinejad Hossein, Takamasa Shirai, Vijay John, Seiichi Mita. Applicant/Owner: SOKEN Co., Ltd., DENSO Corporation, Toyota School Foundation. *Patent No. JP6701057B2*, Japan. Filed: November 4, 2016. Published: May 10, 2018. Application No.: 2016-215759. Publication No.: 2018-073308.

PUBLICATIONS

Google Scholar: Vijay John

Book chapter

1. N. Ramesh, P. Joe, and V. John, "Surgical Robots in Urology and Gynecology", Chapter 4 in *Surgical Robots in Smart Hospitals*, 2025, <https://doi.org/10.1002/9781394355358.ch4>.
2. Y. Jawale, P. Joe, and V. John, "Next-Generation Industrial Robots", Chapter 21 in *Surgical Robots in Smart Hospitals*, 2025, <https://doi.org/10.1002/9781394355358.ch21>
3. L. Shuo, V. John and Z. Liu, "On the Prospects of Using Deep Learning for Surveillance and Security Applications", *Deep Learning for Image Processing Applications, Advances in Parallel Computing*, 2017, DOI 10.3233/978-1-61499-822-8-218
4. V. John, "Human Gait Signature for Biometric Authentication", In *the book, Advances in Biometrics for Secure Human Authentication and Recognition*, CRC Press, Taylor and Francis, ISBN-9781466582422, pp. 450, August 2013.

5. V. John, S. Ivekovic and E. Trucco, “Markerless Human Motion Capture using Hierarchical Particle Swarm Optimisation”, *Chapter in Computer Vision, Imaging and Computer Graphics: Theory and Applications, Communications in Computer and Information Science*, volume 68, 2010, ISBN: 9783642118395

Select Journal Publications

1. T. Nguyen, Y. Kawanishi, V. John, T. Komamizu, and I. Ide, “MultiSensor-Home: Multi-modal multi-view dataset and benchmarks for action recognition in home environments”, in *Pattern Recognition*, Apr 2026, Volume 179, <https://doi.org/10.1016/j.patcog.2026.113810>
2. V. John and Y. Kawanishi, “Hierarchical graph attention networks with spatio-temporal class tokens for distributed audio-visual event classification”, *Multimedia Tools and Applications* Apr 2026, Volume 85, issue 4, <https://doi.org/10.1007/s11042-026-21512-2>
3. T. Nguyen, Y. Kawanishi, V. John, T. Komamizu, and I. Ide, “Action Selection Learning for Weakly Labeled Multi-modal Multi-view Action Recognition”, in *ACM Transactions on Multimedia Computing, Communications, and Applications*, June 2025, doi/10.1145/3744742
4. L. Zhang, Q. Liang, V. John, H. Chen, S. Li, W. Li, Y. Chen, “Intelligent Psyllid Monitoring Based on DiTs-YOLOv10-SOD”, in *IEEE Transactions on AgriFood Electronics*, Apr 2025, Volume: 3, Issue: 1, pages 286-294
5. V. John, Y. Kawanishi, “Multimodal Cascaded Framework with Multimodal Latent Loss Functions Robust to Missing Modalities”. *ACM Transactions of Multimedia Computing, Communication, and Applications*, Jan 2025, 10.1145/3711860
6. C. Yang, Z. Tian, X. You, K. Jia, T. Liu, Z. Pan, V. John, “Polylanenet++: enhancing the polynomial regression lane detection based on spatio-temporal fusion”. *Signal, Image and Video Processing*, 2024, 18(4), 3021-3030
7. V. John, Y. Kawanishi, “Progressive Learning of a Multimodal Classifier Accounting for Different Modality Combinations”. *Sensors*, 2023, 23(10), 4666-4676
8. Z. Bao, W. Li, J. Chen, H. Chen, V. John, C. Xiao, Y. Chen, “Predicting and Visualizing Citrus Color Transformation Using a Deep Mask-Guided Generative Network”, *Plant Phenomics*, 2023, 5, 0057.
9. Xuebin Zhao, Weifu Li, Hong Chen, Yingjie Wang, Yanhong Chen and Vijay John, “Distribution dependent feature selection for deep neural networks”, *Applied Intelligence*, 2022, 52(4), 4432-4442
10. V. John, A. Lakshmanan, S. Mita, A. Boyali and S. Thompson, “Deep Visible and Thermal Camera-based Optimal Semantic Segmentation using Semantic Forecasting”, *ASME Journal of Autonomous Vehicles and Systems*, 2021, 1(2), 021006-021015
11. V. John and S. Mita, “Deep Feature-Level Sensor Fusion Using Skip Connections for Real-Time Object Detection in Autonomous Driving”, *Electronics*, 2021, 10(4), 424
12. Liu, Shuo, Gao, Mingliang, John, Vijay, Liu, Zheng, Blasch, Erik, “Deep Learning Thermal Image Translation for Night Vision Perception”, in *IEEE Transactions of Intelligent Systems and Technology*, Dec 2021, 12(9), 1 - 18.
13. S. Liu, H. Liu, V. John, Z. Liu, E. Blasch and Y. Huang, “Enhanced Situation Awareness through CNN-based Deep Multi-Modal Image Fusion”, in *Optical Engineering*, May 2020, 59(5), 053103
14. M. Gao, J. Jiang, G. Zou, V. John, and Z. Liu, “RGB-D-based Object Recognition using Multimodal Convolutional Neural Networks: A Survey”, in *IEEE Access*, Mar 2019, vol. 7, pp. 43110-43136
15. V. John, Y. Xu and S. Mita, “Stereo Vision based Vehicle Localization in Point Cloud Maps Using Multi-Swarm Particle Swarm Optimization”, in *Signal Image and Video Processing*, Jan 2019, vol. 13, pp. 805-812

16. V. John, A. Boyali, H. Tehrani, K. Ishimaru, M. Konishi, Z. Liu, and S. Mita, "Estimation of Steering Angle and Collision Avoidance for Automated Driving using Deep Mixture of Experts", in *IEEE Transactions on Intelligent Vehicles*, June 2018, vol. 3, no. 4, pp. 571-584
17. V. John, Z. Liu, C. Guo, S. Mita, K. Kidono, H. Tehrani and K. Ishimaru, "Real-time Road Surface and Semantic Lane Estimation using Deep Features", in *Signal, Image and Video Processing*, March 2018, 12(4), pp. 1133-1140
18. Z. Liu, E. Blasch, G. Bhatnagar, V. John, W. Wue, R. Blum, "Fusing Synergistic Information from Multi-Sensor Images: from Implementation to Performance Assessment", in *Information Fusion*, Nov 2017, vol. 42, pp. 127-145
19. V. John, A. Boyali and S. Mita, "Gabor Filter and Gershgorin Disk-based Convolutional Filter Constraining for Image Classification", in *International Journal of Machine Learning and Computing*, Volume 1, Issue 4, pp 55-60, 2017.
20. S. S. Rathour, Ali Boyali, Lyu Zheming, Seichi Mita, and Vijay John, "A Map-based Lateral and Longitudinal DGPS/DR Bias Estimation Method for Autonomous Driving," in *International Journal of Machine Learning and Computing* vol. 7, no. 4, pp. 67-71, 2017.
21. Z. Liu, E. Blasch and V. John, "Statistical Comparison of Pixel-Level Fusion Algorithms", in *Information Fusion*, 2017, vol. 36, pp. 251-260
22. V. John, S. Tsuchizawa, Z. Liu and S. Mita, "Fusion of Thermal and Visible Cameras for the Application of Pedestrian Detection", in *Signal Image and Video Processing*, 2017, vol. 11, pp. 517-524
23. V. John, Q. Long, Y. Xu, Z. Liu and S. Mita, "Sensor Fusion of Lidar and Stereo Camera without Calibration Objects", in *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, Feb. 2017, Vol.E100-A, No.2 [**Invited Paper**].
24. B. Qi, V. John, Z. Liu and S. Mita, "Pedestrian Detection from Thermal Images: A Sparse Representation Based Approach", in *Infrared, Physics and Technology*, Feb 2016, vol. 76, pp 157-167
25. V. John, K. Yoneda, Z. Liu and S. Mita, "Saliency Map Generation by Convolutional Neural Network for Real-time Traffic Light Detection", in *IEEE Transactions on Computational Imaging*, Sept 2015, 1(3), pp. 159-173.
26. Q. Xie, C. Ma, C. Guo, V. John, S. Mita, Q. Long, "Image Fusion Based on the $\Delta^{-1}TV^0$ Energy Function", in *Entropy*, Nov 2014, vol. 16, pp. 6099-6115.
27. V. John and E. Trucco, "Charting-based subspace learning for video-based human action classification", in *Machine Vision and Application*, Special Issue: Multimedia Event Detection, January 2014, 25(1), pp 119-132.
28. V. John, E. Trucco and S. Ivezic, "Markerless human articulated tracking using hierarchical particle swarm optimisation", in *Image Vision Computing*, November 2010, 28(11), pp 1530-1547.

Select Peer-reviewed Conferences

1. T. Nguyen, Y. Kawanishi, V. John, T. Komamizu, I. Ide, "View-aware Cross-modal Distillation for Multi-view Action Recognition", *IEEE/CVF Winter Conference on Applications of Computer Vision*, 2026.
2. E. Malathy, V. John, S. Selcia, S. Vikraman, "Bidirectional Translation System for Tamil Sign Language Recognition", in *Proceedings of Intelligent Vision and Computing*, 2025.
3. Y. Fang, V. John, Y. Kawanishi, "SemGest: A Multimodal Feature Space Alignment and Fusion Framework for Semantic-aware Co-speech Gesture Generation", in *Proceedings of the International Workshop on Generation and Evaluation of Non-verbal Behaviour for Embodied Agents*, 2025

4. Z. Zhu, Z. Jin, J. Zhang, F. Chen, J. Zhou, V. John, F. Spiess, “Reproducibility Companion Paper: Enhancing Model Interpretability with Local Attribution over Global Exploration”, in the Proceedings of ACM International Conference on Multimedia, 2025.
5. T. Nguyen, Y. Kawanishi, V. John, T. Komamizu, I. Ide, “MultiSensor-Home: A Wide-area Multi-modal Multi-view Dataset for Action Recognition and Transformer-based Sensor Fusion”, in Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, 2025 [**best student paper**]
6. J. Y. Chen, V. John, Y. Kawanishi, “Cross-modal Emotion-specific Attention Model for Multimodal Emotion Recognition”, in Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, 2025
7. V. John, Y. Kawanishi, “Modelling Spatio-Temporal Dynamics by Graph Attention Network for Distributed Multi-Microphone Sound Event Classification” in Proceedings of the IEEE International Conference on Advanced Visual and Signal-Based Systems, 2025
8. V. John, Y. Kawanishi, “Generating Pseudo-Strong Labels from Weak Labels for Multi-Source Sound Event Detection”, in Proceedings of the International Conference on Pattern Recognition, 2024
9. V. John, Y. Kawanishi, “Frame-Level Latent Embedding using Weak Labels for Multi-view Action Recognition”, in Proceedings of the IEEE International Conference on Multimedia Information Processing and Retrieval, 2024
10. V. John, Y. Kawanishi, “Combining Knowledge Distillation and Transfer Learning for Sensor Fusion in Visible and Thermal Camera-based Person Classification”, in Proceedings of the International Conference on Machine Vision and Applications, 2023
11. V. John, Y. Kawanishi, “Multimodal Cascaded Framework with Metric Learning Robust to Missing Modalities for Person Classification”, in Proceedings of the 14th Conference on ACM Multimedia Systems, 257-265, 2023
12. V. John, Y. Kawanishi, “Audio-Visual Sensor Fusion Framework using Person Attributes Robust to Missing Visual Modality for Person Recognition”, in Proceedings of the International Conference on Multimedia Modelling, 2023
13. V. John, Y. Kawanishi, “A Multimodal Sensor Fusion Framework Robust to Missing Modalities for Person Recognition”, in Proceedings of ACM Multimedia Asia, 2022
14. V. John, Y. Kawanishi, “Audio and Video-based Emotion Recognition using Multimodal Transformers”, in Proceedings of the International Conference on Pattern Recognition, 2022
15. V. John, A. Lakshmanan, S. Mita, A. Boyali and S. Thompson, “Deep Fusion-based Visible and Thermal Camera Forecasting using Seq2Seq GAN”, in Proceedings of the IEEE Intelligent Vehicle Symposium, 2021
16. W. Li, V. John, S. Mita, “Enhancing Depth Quality of Stereo Vision using Deep Learning-based Prior Information of the Driving Environment”, in Proceedings of the International Conference on Pattern Recognition, 2021
17. V. John, S. Mita, A. Boyali and S. Thompson, “BVTNet: Multi-label Multi-class Fusion of Visible and Thermal Camera for Free Space and Pedestrian Segmentation”, in Proceedings of the International Conference on Pattern Recognition Workshop, 2021
18. V. John, S. Mita, A. Boyali and S. Thompson, “Visible and Thermal Camera-based Jaywalking Estimation using a Hierarchical Deep Learning Framework”, in Proceedings of the Asian Conference Computer Vision Workshop, 2020
19. V. John, S. Mita, “RVNet: Deep Sensor Fusion of Monocular Camera and Radar for Real-Time Obstacle Detection in Challenging Environments”, in the Proceedings of the Pacific Rim Symposium on Image and Video Technology, 2019

20. V. John, M. K. Nithilan, S. Mita, H. Tehrani, R. S. Sudheesh, and P. P. Lalu, "SO-Net: Joint Semantic Segmentation and Obstacle Detection using Deep Fusion of Monocular Camera and Radar", in the Proceedings of the Pacific Rim Symposium on Image and Video Technology (Workshops), 2019.
21. A. Boyali, T. Akarman, N. Hashimoto, V. John, "Multi-Agent Reinforcement Learning for Autonomous On Demand Vehicles", in Proceedings of the IEEE Intelligent Vehicle Symposium, 2019
22. V. John, N. Karunakaran, C. Guo, K. Kidono, S. Mita, "Free Space, Visible and Missing Lane Marker Estimation using the PsiNet and Extra Trees Regression", in Proceedings of the International Conference on Pattern Recognition, 2018
23. V. John, N. Karunakaran, S. Mita, H. Tehrani, M. Konishi, K. Ishimaru, Y. Xu, "Sensor Fusion of Intensity and Depth Cues using the ChiNet for Semantic Segmentation of Road Scenes", in Proceedings of the IEEE Intelligent Vehicle Symposium, 2018
24. S. Rathour, V. John, N. Karunakaran, and S. Mita, "Vision and Dead Reckoning-based End-to-End Parking for Autonomous Vehicles", in Proceedings of the IEEE Intelligent Vehicle Symposium, 2018
25. V. John, S. Rathour, S. Mita, H. Tehrani, K. Ishimaru, M. Konishi, and H. Chin, "Vision-based Steering Angle Prediction by the Fusion of Depth and Intensity Deep Features", in Proceedings of the Proceedings of the Indian Conference on Computer Vision, Graphics and Image Processing, 2018.
26. A. Boyali, L. Zheming, V. John, S. Mita and T. Yoshikawa, "Self-Scheduling Robust Preview Controllers for Path Tracking and Autonomous Vehicles", in Proceedings of Intelligent Transportation Systems Conference, 2017
27. V. John, Q. Long, Y. Xu and S. Mita, "Registration of GPS and Stereo Vision for Point Cloud Localization in Intelligent Vehicles Using Particle Swarm Optimization", in Proceedings of International Conference of Swarm Intelligence, 2017
28. V. John, M. Umetsu, A. Boyali, S. Mita, N. Imanashi and S. Sanma, "Real-time Hand Posture and Gesture-based Touchless Automotive User Interface using Deep Learning", in Proceedings of the IEEE Intelligent Vehicle Symposium, 2017
29. V. John, S. Mita, H. Tehrani and K. Kidono, "Automated Driving by Monocular Camera Using Deep Mixture of Experts", in Proceedings of the IEEE Intelligent Vehicle Symposium, 2017 **[Invited for IEEE Transactions]**
30. V. John, A. Boyali, S. Mita, M. Imanishi and N. Sanma, "Deep Learning-based Fast Hand Gesture Recognition using Representative Frames", in Proceedings of International Conference on Digital Image Computing: Techniques and Applications, 2016.
31. V. John, C. Guo, S. Mita, K. Kidono, H. Tehrani and K. Ishimaru, "Fast Road Scene Segmentation using Deep Learning and Scene-based Models", in Proceedings of the International Conference on Pattern Recognition, 2016.
32. V. John, L. Zheming and S. Mita, "Robust Traffic Light and Arrow Detection using Optimal Camera Parameters and GPS-based Prior", in the Proceedings of the IEEE Asia-Pacific Conference on Intelligent Robot Systems, 2016. **[Best Image Processing Oral Presentation]**
33. V. John, Z. Liu, C. Guo, S. Mita and K. Kidono, "Real-Time Lane Estimation using Deep Features and Extra Trees Regression", in the Proceedings of the Pacific Rim Symposium on Image and Video Technology, 2015. **[Best Applications Paper]**
34. V. John, Q. Long, Z. Liu and S. Mita, "Automatic Calibration and Registration of Lidar and Stereo Camera without Calibration Objects", in the Proceedings of the IEEE International Conference on Vehicular Electronics and Safety, 2015.

35. V. John, Z. Liu, S. Mita and B. Qi, "Pedestrian Detection in Thermal Images Using Adaptive Fuzzy C-Means Clustering and Convolutional Neural Networks", in the Proceedings of the International Conference on Machine Vision Application, 2015.
36. V. John, K. Yoneda, Z. Liu, S. Mita and B. Qi, "Traffic Light Recognition in Varying Illumination using Deep Learning and Saliency Map", in the Proceedings of the IEEE Conference on Intelligent Transportation Systems, 2014.
37. B. Qi, V. John, Z. Liu, and S. Mita, "Pedestrian Detection from Thermal Images with A Scattered Difference of Directional Gradients Feature Descriptor, in the Proceedings of the IEEE Conference on Intelligent Transportation Systems, 2014.
38. B. Qi, V. John, Z. Liu, and S. Mita, "Use of Sparse Representation for Pedestrian Detection in Thermal Images", in the Proceedings of Computer Vision and Pattern Recognition Workshops, 2014
39. V. John, G. Englebienne and B. Krose, "Solving Person Re-identification using Efficient Gibbs Sampling", in the Proceedings of British Machine Vision Conference, 2013. [**Oral 7% acceptance rate**]
40. V. John, G. Englebienne and B. Krose, "Person Re-identification using Height-based Gait in Colour Depth Cameras", in the Proceedings of International Conference of Image Processing, 2013.
41. V. John, G. Englebienne and B. Krose, "Relative Camera Localisation in Non-Overlapping Camera Networks using Multiple Trajectories", in the Proceedings of the European Conference of Computer Vision Workshops, 2012
42. V. John, E. Trucco and S. J. McKenna, "Markerless human motion capture using charting and manifold constrained particle swarm optimisation", in the Proceedings of the BMVC UK postgrad.workshop, 2010. [**Best Paper**]
43. S. Ivekovic, V. John, and E. Trucco, "Markerless Multi-view Articulated Pose Estimation Using Adaptive Hierarchical Particle Swarm Optimisation", in the Proceedings of the Applications of Evolutionary Computation, 2010
44. V. John and E. Trucco, "Multiple view human articulated tracking using charting and particle swarm optimisation", in the Proceedings of the ACM Multimedia International workshop on 3D Video Processing, 3DVP, 2010.
45. V. John, S. Ivekovic and E. Trucco, "Articulated Human Motion Tracking with HPSO", in the Proceedings of VISSAPP, 2009.

UNIVERSITY ADMINISTRATION

- Chair of Vision Centric Challenge, Robofest 2026
- Executive Board Member, Robofest.
- Member of *I-Plaza* Committee at TTI, Japan. Responsibilities include language training for Japanese students, and integration of foreign researchers into the university environment.
- Leading the efforts in establishing a collaborative tie-up between the Nodal Center for Robotics and Artificial Intelligence, India and Research Center for Smart Vehicles, Japan.
- Organized and Delivered a 2-Day workshop on Computer Vision, Image Processing and Machine Learning, Karunya University, India.

CONTRIBUTION TO THE RESEARCH COMMUNITY

Program Committee

IEEE International Conference on Intelligent Transportation Systems, IEEE Intelligent Vehicle Symposium, International Conference on Data Intelligence and Security, International Indian Conference on Computer Vision, Graphics and Image Processing, Multimedia Modelling

Reviewing Duties

I have served as a reviewer for several leading journals and conferences, including *IEEE Access*, *IEEE Transactions on Emerging Topics in Computational Intelligence*, *IEEE Transactions on Affective Computing*, *IEEE Transactions on Industrial Informatics*, *IEEE Transactions on Instrumentation & Measurement*, *IEEE Transactions of Fuzzy Systems*, *IEEE Transactions of Intelligent Vehicles*, *IEEE Transactions of Intelligent Transportation Systems*, *IEEE Transactions on Systems, Man and Cybernetics: Systems*, *Computer Vision and Image Understanding*, *IEEE Transactions of Image Processing*, *Journal of Information Fusion*, *Robotics and Autonomous Systems*, *Pattern Analysis and Applications*, *Signal and Image Video Processing*, *Journal of Ambient Intelligence and Humanized Computing*, *Frontiers in Neuroanatomy*, *International Journal of Information Technology & Decision Making*, *Sensors*, *Machine Vision Applications*, *Infrared Physics and Technology*, and *ACM Transactions on Intelligent Systems and Technology*.

INVITED LECTURES

- “Missing Modality Problem for Perception,” ICCCS 2023, India
- “Multimodal Learning for Perception in Intelligent Mobility,” Workshop, ICCCS 2023, India
- “Deep Sensor Fusion for Thermal and Stereo Cameras,” GTC 2019, USA
- “Deep Sensor Fusion for Stereo Cameras,” GTC 2018, Japan
- “Environment Perception for Autonomous Driving,” IV Workshop, 2018
- “Deep Learning for Autonomous Driving,” Karunya University, India, 2018