

Anoop Cherian, Ph.D.

Senior Principal Research Scientist, Mitsubishi Electric Research Labs (MERL)

Adjunct Associate Professor, Australian National University

201 Broadway, 8th Floor, Cambridge, MA 02139, USA

Email: firstname.lastname@gmail.com | Website: <https://users.cecs.anu.edu.au/~cherian>

Research Interests

Towards building **multimodal physically-grounded world models for AI reasoning** combining multimodal foundation models, embodied multi-agent behavior modeling and orchestration, physics-grounded imagination, and diffusion-based generative approaches with applications in multi-agent embodied systems and robotics (see publications).

Education

- | | |
|-----------|--|
| 2010–2013 | Ph.D. Computer Science, University of Minnesota, Minneapolis |
| 2007–2010 | M.S. Computer Science, University of Minnesota, Minneapolis |
| 1998–2002 | B.Tech. (Hons.), Computer Science & Engineering, NIT Calicut/Kerala, India |

Current Appointment

- **Senior Principal Research Scientist**, MERL, Cambridge MA (2023–present)
- **Adjunct Senior Lecturer**, Australian National University, Canberra (2015–present)

Previous Positions

- Principal / Research Scientist, MERL, 2017–2023
- Research Fellow, Australian Centre for Robotic Vision, ANU, 2015–2017
- Post-doctoral Researcher, INRIA Rhône-Alpes (LEAR), 2012–2015
- Research Intern, Max Planck Institute, Tübingen, 2009
- Software Design Engineer, Microsoft, 2005–2007

At-a-Glance Metrics

Google Scholar *h*-index: 41+; *i10*-index: 88+ Top-tier pubs: 16×CVPR, 7×ICCV, 5×ICML, 4×NeurIPS, 4×ECCV, 5×AAAI, 7×TPAMI

Selected Honors & Awards

- **Area Chair:** NeurIPS (2023–26), ICML (2023–25), ECCV 24, ICLR 25, AAAI 2023–24
- **Keynote speaker:** AGI 2023 (“Are Deep Neural Networks SMARTer than Second Graders?”)
- Best Paper Runners Up, SPaVLE Workshop @ NeurIPS’25 for SpinBench paper.
- MERL President’s Award (2025), Multimodal Algorithmic Reasoning
- MELCO Director’s Award (2023); MERL President’s Awards (2020–23)
- Outstanding Reviewer: ICML 2022, NeurIPS 2020, CVPR 2017, ECCV 2016
- Best Student Paper, ICIP 2012; Best Poster, Minnesota Supercomputing Institute 2012

Scientific Leadership (selected)

- Lead organizer, Multimodal Algorithmic Reasoning Workshops (NeurIPS-2024-25, CVPR 2024-25)
- Lead organizer, Vision & Language Algorithmic Reasoning Workshop & SMART-101 Challenge (ICCV 2023)

- Creator, SMART-101 and SMART-840 datasets (CVPR 2023, NeurIPS 2024), **AVLEN** benchmark (NeurIPS 2023)

Teaching & Service

- Adjunct Senior Lecturer, ANU: guest lectures, Ph.D. committees (2015–present)
- Teaching Assistant, "Artificial Intelligence", University of Minnesota (2010)
- Guest Editor, *Journal of Imaging* (Special Issue on Action Recognition)
- Reviewer/PC member: CVPR, ICCV, ECCV, ICML, NeurIPS, AAAI; journals incl. *IEEE TPAMI, IJCV, TNNLS, TIP, JMLR*

Student Supervision (highlights)

- Jue Wang (Ph.D. ANU, 2016–21): 3 TPAMI, CVPR, ECCV (first appointment: Facebook AI)
- Jiahao Zhang (Ph.D. ANU, 2021–): CVPR, WACV, ICCV (first Appointment, Roblox)
- Anshul Shah (Ph.D. UMD, 2020-2022): AAAI, CVPR, NeurIPS (first appointment: Apple)
- Moitreya Chatterjee (Ph.D. UIUC, 2019–22): NeurIPS, ECCV, ICCV (first appointment: MERL)
- Sudipta Paul (Ph.D. UCR, 2021–22): NeurIPS 2022 (first appointment: Samsung)
- Mentored 30+ students/interns, yielding 60+ joint publications

Full Publications List

* indicates the student I supervised for the project.

Book Chapters

1. **A. Cherian** and S. Sra, "Positive Definite Matrices: Data Representation and Application to Computer Vision", Algorithmic Advances in Riemannian Geometry and Applications - For Machine Learning, Computer Vision, Statistics, and Optimization, Springer, 2016

Conference Articles and Pre-prints

2. Q. Zhou, C. Gan, and **A. Cherian**, LLawCo: Learning Laws of Cooperation for Modeling Embodied Multi-Agent Behavior, ICML, 2026
3. D. Li*, J. Zhang*, B. Egger, M. Chatterjee, S. Lohit, T. K. Marks, and **A. Cherian**, AssemblyBench: Physics-Aware Assembly of Complex Industrial Objects, CVPR, 2026
4. Y. Zhang, R. Corcodel, C. Hori, **A. Cherian**, and D. Zhao, SpinBench: Perspective and Rotation as a Lens on Spatial Reasoning in VLMs, ICLR, 2026
5. **A. Cherian**, R. Corcodel, S. Jain, and D. Romeres, LLMPhy: Parameter Identifiable Physical Reasoning Combining Large Language Models and Physics Engines, AISTATS, 2026
6. K. Kogashi*, **A. Cherian**, and M-Y. Kuo, MMHOI: Modeling Complex 3D Multi-Human Multi-Object Interactions, WACV, 2026 (oral)
7. **A. Cherian**, R. Doyle*, E. Ben-Dov*, S. Lohit, and K-C. Peng, WISE: Weighted Iterative Society-of-Experts for Multimodal Debate with Probabilistic Consensus, abs/2512.02405, 2025
8. J. Zhang*, **A. Cherian**, C. Rodriguez-Opazo, W. Deng, S. Gould, Manual-PA: Learning 3D Part Assembly from Instruction Diagrams, ICCV, 2025
9. **A. Cherian**, R. Corcodel, S. Jain, and D. Romeres, LLMPhy: Complex Physical Reasoning Using Large Language Models and World Models, Preprint, 2025
10. J. Zhang, F. Zhang, C. Rodriguez, Y. Ben-Shabat, **A. Cherian**, and S. Gould, Temporally

Grounding Instructional Diagrams in Unconstrained Videos, WACV, 2025

11. Y. He*, S. Shin, **A. Cherian**, N. Trigoni, and A. Markham, SoundLoc3D: Invisible 3D Sound Source Localization and Classification Using a Multimodal RGB-D Acoustic Camera, WACV 2025
12. **A. Cherian**, K.-C. Peng, S. Lohit, J. Matthiesen, K. Smith, and J. Tenenbaum, Evaluating Large Vision-and-Language Models on Children’s Mathematical Olympiads, NeurIPS, 2024
13. **A. Cherian**, S. Jain, and T. K. Marks, Few-shot Transparent Instance Segmentation for Bin-picking, IROS, 2024
14. J. Yin, A. Luo, Y. Du, **A. Cherian**, T. K. Marks, J. L. Roux, and C. Gan, Disentangled Acoustic Fields For Multimodal Physical Scene Understanding, IROS 2024
15. Y. He*, **A. Cherian**, G. Wichern, and A. Markham, Deep Neural Room Acoustics Primitive, ICML 2024
16. H. Ni, B. Egger, S. Lohit, **A. Cherian**, Y. Wang, T. Koike-Akino, S. X. Huang, and T. Marks, “TI2V-Zero: Zero-Shot Image Conditioning for Text-to-Video Diffusion Models”, Computer Vision and Pattern Recognition (CVPR), 2024
17. Z. Yang, J. Liu, P. Chen, **A. Cherian**, T. K. Marks, J. Le Roux, and C. Gan, “RILA: Reflective and Imaginative Language Agent for Zero-Shot Semantic Audio-Visual Navigation”, Computer Vision and Pattern Recognition (CVPR), 2024
18. X. Liu*, S. Paul, M. Chatterjee, and **A. Cherian**, “CAVEN: An Embodied Conversational Agent for Efficient Audio-Visual Navigation in Noisy Environments”, AAAI, 2024
19. X. Zhu*, D. K. Jha, D. Romeres, L. Sun, M. Tomizuka, and **A. Cherian**, “Multi-level Reasoning for Robotic Assembly: From Sequence Inference to Contact Selection”, International Conference on Robotics and Automation (ICRA), 2024
20. Z. Carmichael, S. Lohit, **A. Cherian**, M. Jones, W. Scheirer, “Pixel-Grounded Prototype Part Networks”, Winter Conference in Computer Vision (WACV), 2024
21. Y. He*, S. Shin, **A. Cherian**, A. Markham, N. Trigoni, “Sound3DVEDet: 3D Sound Source Detection using Multiview Microphone Array and RGB Images”, Winter Conference in Computer Vision (WACV), 2024
22. C. Hori, P. Wang, M. Rahman, C. Vaca-Rubio, S. Khurana, **A. Cherian**, and J. Le Roux, “Wi-Fi based Indoor Monitoring Enhanced by Multimodal Fusion”, International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2024
23. **A. Cherian**, K. Peng, S. Lohit, K. Smith, and J. B. Tenenbaum, “Are Deep Neural Networks SMARTer than Second Graders?”, Computer Vision and Pattern Recognition (CVPR), 2023
24. N. Nair, **A. Cherian**, S. Lohit, Y. Wang, T. Koike-Akino, V. Patel, and T. Marks, “Steered Diffusion: A Generalized Framework for Plug-and-Play Conditional Image Synthesis”, Intl. Conference on Computer Vision (ICCV), 2023
25. J. Zhang, **A. Cherian**, Y. Liu, Y. Ben-Shabat, C. Rodriguez, and S. Gould, “Aligning Step-by-Step Instructional Diagrams to Video Demonstrations”, Computer Vision and Pattern Recognition (CVPR), 2023

26. A. Shah, A. Roy, K. Shah, S. K. Mishra, D. Jacobs, **A. Cherian**, and R. Chellappa, “HaLP: Hallucinating Latent Positives for Skeleton-based Self-Supervised Learning of Actions”, Computer Vision and Pattern Recognition (CVPR), 2023
27. K. Ota, H. Tung, K. Smith, **A. Cherian**, T. K. Marks, A. Sullivan, A. Kanezaki, and J. B. Tenenbaum, “H-SAUR: Hypothesize, Simulate, Act, Update, and Repeat for Understanding Object Articulations from Interactions”, International Conference on Robotics and Automation (ICRA), 2023
28. **A. Cherian**, S. Jain, T. K. Marks, and A. Sullivan, “Discriminative 3D Shape Modeling for Few-shot Instance Segmentation”, International Conference on Robotics and Automation (ICRA), 2023
29. S. Paul*, A. Chowdhary, and **A. Cherian**, “AVLEN: Audio-Visual-Language Embodied Navigation in 3D Environments”, Advances in Neural Information Processing Systems (NeurIPS), 2022
30. M. Chatterjee*, N. Ahuja, and **A. Cherian**, “Learning Audio-Visual Dynamics Using 2.5D Scene Graphs for Audio Source Separation”, Advances in Neural Information Processing Systems (NeurIPS), 2022
31. A. Roy, A. Shah, K. Shah, P. Dhar, **A. Cherian**, and R. Chellappa, “FeLMi: Few shot Learning with hard Mixup”, Advances in Neural Information Processing Systems (NeurIPS), 2022
32. **A. Cherian**, C. Hori, T. Marks, and J. Le Roux, “(2.5+1)D Spatio-Temporal Scene Graphs for Video Question Answering”, AAAI Conference on Artificial Intelligence (AAAI), 2022 (**Oral presentation, 3% acceptance**)
33. A. Shah*, S. Sra, R. Chellappa, and **A. Cherian**, “Max-Margin Contrastive Learning”, AAAI Conference on Artificial Intelligence (AAAI), 2022 (15)
34. S. Medin, B. Egger, **A. Cherian**, Y. Wang, J. Tenenbaum, X. Liu, and T. Marks, “MOST-GAN: 3D Morphable StyleGAN for Disentangled Face Image Manipulation”, AAAI Conference on Artificial Intelligence (AAAI), 2022 (**15% acceptance**)
35. A. Shah, S. Geng, P. Gao, **A. Cherian**, T. Hori, T. Marks, J. Roux, C. Hori, “Audio-Visual Scene-Aware Dialog and Reasoning using Audio-Visual Transformers with Joint Student-Teacher Learning”, International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2022
36. M. Chatterjee*, N. Ahuja, and **A. Cherian**, “A Hierarchical Variational Neural Uncertainty Model for Stochastic Video Prediction”, International Conference on Computer Vision (ICCV), 2021 (**Oral presentation, 3% acceptance**)
37. **A. Cherian**, G. Pais*, S. Jain, T. Marks, and A. Sullivan, “InSeGAN: A Generative Approach to Segmenting Identical Instances in Depth Images”, International Conference on Computer Vision (ICCV), 2021
38. M. Chatterjee*, J. Le Roux, N. Ahuja, and **A. Cherian**, “Visual Scene Graphs for Audio Source Separation”, International Conference on Computer Vision (ICCV), 2021
39. S. Geng*, P. Gao, M. Chatterjee*, C. Hori, J. Le Roux, Y. Zhang , H. Li, and **A. Cherian**,

- “Dynamic Graph Representation Learning for Video Dialog via Multi-Modal Shuffled Transformers”, AAAI Conference on Artificial Intelligence (AAAI), 2021
40. M. Benosman, O. Romero, and **A. Cherian**, “Optimizing Deep Neural Networks via Discretization of Finite-Time Convergent Flows”, arXiv, <https://arxiv.org/abs/2010.02990>, 2021
 41. S. Kumar, L. Van Gool, C. de Oliveira, **A. Cherian**, Y. Dai, and H. Li, “Dense Non-Rigid structure from motion: A manifold viewpoint”, arXiv:2006.09197, 2020
 42. M. Chatterjee* and **A. Cherian**, “Sound2Sight: Generating Visual Dynamics from Sound and Context”, European Conference on Computer Vision (ECCV), 2020
 43. **A. Cherian**, J. Wang*, C. Hori, and T.K. Marks, “Spatio-Temporal Ranked-Attention Networks for Video Captioning”, IEEE Winter Conference on Computer Vision (WACV), 2020
 44. R. Huang, W. Xu, T.Y. Lee, A. Cherian, Y. Wang, and T. Marks, “FX-GAN: Self-supervised GAN learning via feature exchange”, IEEE Winter Conference on Computer Vision (WACV), 2020
 45. A. Kumar, T. K. Marks, W. Mou, Y. Wang, M. Jones, **A. Cherian**, T. Koike-Akino, X. Liu, and C. Feng, “LUVLi Face Alignment: Estimating Landmarks’ Location, Uncertainty, and Visibility Likelihood”, Computer Vision and Pattern Recognition (CVPR), 2020
 46. A. Raghunathan, **A. Cherian**, and D. Jha, “Game Theoretic Optimization via the Gradient-based Nikaido-Isoda Function”, International Conference on Machine Learning (ICML), 2019
 47. H. Alamri, V. Cartillier, A. Das, J. Wang, S. Lee, P. Anderson, I. Essa, D. Parikh, D. Batra, **A. Cherian**, T. Marks, and C. Hori, “Audio-Visual Scene-Aware Dialog”, Computer Vision and Pattern Recognition (CVPR), 2019
 48. C Hori, A Cherian, TK Marks, and T Hori, “Joint Student-Teacher Learning for Audio-Visual Scene-Aware Dialog.”, INTERSPEECH, 2019
 49. J. Wang* and **A. Cherian**, GODS: “Generalized One-class Discriminative Subspaces for Anomaly Detection”, International Conference on Computer Vision (ICCV), 2019
 50. C. Hori, H. Alamri, J. Wang, G. Wichern, T. Hori, **A. Cherian**, T. K Marks, V. Cartillier, R.Gontijo Lopes, A. Das, I. Essa, D. Batra, and D. Parikh, “End-to-end audio visual scene-aware dialog using multimodal attention-based video features”, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019
 51. **A. Cherian** and A. Sullivan, “Sem-GAN: Semantically-consistent Image-to-Image Translation”, IEEE Winter Conference on Applications of Computer Vision (WACV), 2019
 52. J. Wang* and **A. Cherian**, “Discriminative Video Representation Learning Using Adversarial Perturbations”, European Conference on Computer Vision (ECCV), 2018 (Oral presentation, acceptance rate <2.5
 53. **A. Cherian**, S. Sra, S. Gould, and R. Hartley, “Non-Linear Temporal Subspace Representations for Activity Recognition”, Computer Vision and Pattern Recognition (CVPR), 2018
 54. S. Kumar*, **A. Cherian**, Y. Dai, and H. Li, “Scalable Non-rigid Structure from Motion: A Grassmannian Perspective”, Computer Vision and Pattern Recognition (CVPR), 2018
 55. J. Wang*, **A. Cherian**, F. Porikli, and S. Gould, “Video Representation Learning Using Discriminative Pooling”, Computer Vision and Pattern Recognition (CVPR), 2018

56. R. Santa Cruz*, B. Fernando, **A. Cherian**, and S. Gould, “Neural Algebra of Classifiers”, Winter Conference on Applications of Computer Vision (WACV), 2018
57. S. Toyer*, **A. Cherian**, T. Han*, and S. Gould, “Human Pose Forecasting via Deep Markov Models”, Intl. Conf. on Digital Image Computing: Techniques and Applications (DICTA), 2017
58. **A. Cherian**, B. Fernando, M. Harandi, and S. Gould, “Generalized Rank Pooling for Action Recognition”, Computer Vision and Pattern Recognition (CVPR), 2017
59. R. SantaCruz*, B. Fernando, **A. Cherian**, and S. Gould, “DeepPermNet: Visual Permutation Learning”, Computer Vision and Pattern Recognition (CVPR), 2017
60. **A. Cherian**, P. Stanitsas*, M. Harandi, V. Morellas, and N. Papanikolopoulos, “Learning Discriminative Alpha-Beta Divergences for Positive Definite Matrices”, International Conference on Computer Vision (ICCV), 2017
61. **A. Cherian**, P. Koniusz, and S. Gould, “Higher-order Pooling on CNN Features via Kernel Linearization for Action Recognition”, Winter Conference on Applications of Computer Vision (WACV), 2017
62. J. Wang*, **A. Cherian**, and F. Porikli, “Ordered Pooling of Optical Flow Sequences for Action Recognition”, Winter Conference on Applications of Computer Vision (WACV), 2017
63. P. Stanitsas*, **A. Cherian**, A. Truskinovsky, V. Morellas, and N. Papanikolopoulos, “Active Convolutional Neural Networks for Cancerous Tissue Recognition”, IEEE Conference on Image Processing (ICIP), 2017 (oral presentation)
64. C. Peng*, H. Li, **A. Cherian**, and H. Yao, “Part-based Fine-grained Bird Image Retrieval Respecting Species Correlation”, IEEE International Conference on Image Processing, (ICIP), 2017
65. S. Gould, B. Fernando, **A. Cherian**, P. Anderson, R. Santa Cruz, and E. Guo, “On Differentiating Parameterized Argmin and Argmax with Application to Bi-level Optimization”, CoRR arxiv:1607.05447, 2016
66. P. Koniusz, **A. Cherian**, and F. Porikli, “Tensor Representations Using Kernel Linearization for Action Recognition from 3D Skeletons”, European Conference on Computer Vision (ECCV), 2016, Springer
67. P. Koniusz and **A. Cherian**, “Sparse Coding for Third-order Supersymmetric Tensors with Application to Texture Recognition”, Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2016, IEEE (Spotlight paper, Top 10)
68. Y. Wen, N. Zhong, X. Yang, and **A. Cherian**, “Riemannian Sparse Coding for Classification of PolSAR Images”, IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2016, IEEE (Oral Presentation)
69. P. Stanitsas*, **A. Cherian**, V. Morellas, and N. Papanikolopoulos, “Active Constrained Clustering via Non-Iterative Uncertainty Sampling”, Proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS), 2016, IEEE
70. P. Stanitsas*, **A. Cherian**, X. Li, A. Truskinovsky, V. Morellas, and N. Papanikolopoulos, “Evaluation of Feature Descriptors for Cancerous Tissue Recognition”, IEEE International Conference on Pattern Recognition (ICPR), 2016, IEEE (Oral Presentation)

71. **A. Cherian**, J. Mairal, K. Alahari, and C. Schmid, “Mixing Body-part Sequences for Human Pose Estimation”, Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2014
72. **A. Cherian**, “Nearest Neighbors Using Compact Sparse Codes”, Proceedings of the 31st International Conference on Machine Learning (ICML), 1053—1061, 2014
73. **A. Cherian** and S. Sra, “Riemannian Sparse coding of Positive Definite Matrices”, Proceedings of the European Conference on Computer Vision (ECCV), 299—314, 2014, Springer International Publishing
74. **A. Cherian**, V. Morellas, and N. Papanikolopoulos, “Robust Sparse Hashing”, Proceedings of the International Conference on Image Processing (ICIP), 2417—2421, 2012, IEEE (**Best Student Paper Award**)
75. R. Sivalingam, **A. Cherian**, V. Morellas, N. Papanikolopoulos, G. Sapiro, et al., “A Multi-Sensor Visual Tracking System for Behavior Monitoring of At-Risk Children”, Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), 1345—1350, 2012, IEEE
76. D. Fehr, **A. Cherian**, R. Sivalingam, S. Nikolay, V. Morellas, and N. Papanikolopoulos, “Compact Covariance Descriptors in 3D Point Clouds for Object Recognition”, Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), 1793—1798, 2012, IEEE
77. **A. Cherian**, S. Sra, A. Banerjee, and N. Papanikolopoulos, “Efficient Similarity Search for Covariance Matrices via the Jensen-Bregman Logdet Divergence”, Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2399—2406, 2011, IEEE (Spotlight Presentation)
78. **A. Cherian**, V. Morellas, N. Papanikolopoulos, and S. Bedros, “Dirichlet Process Mixture Models on Symmetric Positive Definite Matrices for Appearance Clustering in Video Surveillance Applications”, Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 3417—3424, 2011, IEEE
79. S. Sra and **A. Cherian**, “Generalized Dictionary Learning for Symmetric Positive Definite Matrices with Application to Nearest Neighbor Retrieval”, Machine Learning and Knowledge Discovery in Database (ECML), 318—332, 2011, Springer Berlin Heidelberg
80. **A. Cherian**, S. Sra, and N. Papanikolopoulos, “Denoising Sparse Noise via Online Dictionary Learning”, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2060—2063, 2011, IEEE
81. **A. Cherian**, V. Morellas, and N. Papanikolopoulos, “Approximate Nearest Neighbors via Dictionary Learning”, SPIE Defense, Security, and Sensing, 80581M-80581M, 2011, International Society for Optics and Photonics (**Oral Presentation**)
82. **A. Cherian**, V. Morellas and N. Papanikolopoulos, “Efficient Similarity Search via Sparse Coding”, Technical Report, Department of Computer Science and Engineering, University of Minnesota, 2011
83. **A. Cherian**, J. Andersh, V. Morellas, B. Mettler and N. Papanikolopoulos, “Motion Estimation of a Miniature Helicopter using a Single Onboard Camera”, American Control Conference (ACC), 4456—4461, 2010, IEEE

84. **A. Cherian**, J. Andersh, V. Morellas, N. Papanikolopoulos, and B. Mettler, “Autonomous Altitude Estimation of a UAV using a Single Onboard Camera”, Proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS), 3900—3905, 2009, IEEE
85. **A. Cherian**, V. Morellas, and N. Papanikolopoulos, “Accurate 3D Ground Plane Estimation from a Single Image”, Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), 2243—2249, 2009, IEEE

Journal Articles

86. F. Mumcu, M. J. Jones, **A. Cherian**, and Y. Yilmaz, Leveraging Multimodal LLM Descriptions of Activity for Explainable Semi-Supervised Video Anomaly Detection, TMLR, 2026
87. C. Peng*, **A. Cherian**, R. Lan, X. Luo, and H. Yao, “Searching by parts: Towards fine-grained image retrieval respecting species correlation”, Gene Expression Patterns, 2023
88. **A. Cherian** and J. Wang*, “Generalized One-class Learning Using Pairs of Complementary Classifiers”, IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2021
89. **A. Cherian**, P. Stanitsas*, J. Wang, M. Harandi, V. Morellas, and N. Papanikolopoulos, “Learning Log-Det Divergences for Positive Definite Matrices”, IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2021
90. S. Kim, M. Galley, C. Gunasekara, S. Lee, A. Atkinson, B. Peng, H. Schulz, J. Gao, J. Li, M. Adada, M. Huang, L. Lastras, J. K Kummerfeld, W. Lasecki, C. Hori, **A. Cherian**, T. Marks, A. Rastogi, X. Zang, S. Sunkara, R. Gupta, “Overview of the eighth dialog system technology challenge: DSTC8”, IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP), 2021
91. P. Koniusz, L. Wang, and **A. Cherian**, “Tensor Representations for Action Recognition”, IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2021
92. P. Stantisas*, **A. Cherian**, A. Turkinovsky, V. Morellas, and N. Papanikolopoulos, “Image Descriptors for Weakly Supervised Histopathological Breast Cancer Data”, Frontiers in Digital Health, 2020
93. J. Wang* and **A. Cherian***, “Discriminative Video Representation Learning Using Support Vector Classifiers”, IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2020
94. **A. Cherian** and S. Gould, “Second-order Temporal Pooling for Action Recognition”, International Journal of Computer Vision (IJCV), 2018
95. R. Santa Cruz*, B. Fernando, **A. Cherian**, and S. Gould, “Visual Permutation Learning”, IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2018
96. N. Zhong, W. Yang, **A. Cherian**, and X. Yang, “Unsupervised Classification of Polarimetric SAR Images via Riemannian Sparse Coding”, IEEE Transactions on Geoscience and Remote Sensing, 2017
97. **A. Cherian** and S. Sra, “Riemannian Dictionary Learning and Sparse Coding for Positive Definite Matrices”, IEEE Transactions on Neural Networks and Learning Systems (TNNLS), 2016 (IF: 4.84)
98. **A. Cherian**, V. Morellas, and N. Papanikolopoulos, “Bayesian Nonparametric Clustering of

Positive Definite Tensors”, IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2015

99. J. Andersh, **A. Cherian**, B. Mettler, and N. Papanikolopoulos, “A Vision based Ensemble Approach to Velocity Estimation for Miniature Rotorcraft”, *Autonomous Robots*, 39(2), 123–138, 2015
100. **A. Cherian**, V. Morellas, and N. Papanikolopoulos, “Efficient Similarity Search via Robust Sparse Hashing”, *IEEE Transactions on Image Processing*, 23(8), 3646–3655, 2014
101. **A. Cherian**, S. Sra, A. Banerjee, and N. Papanikolopoulos, “Jensen-Bregman Logdet Divergence for Efficient Distance Computations on Positive Definite Matrices”, *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 35(9), 2161–2174, 2013
102. G. Somasundaram, **A. Cherian**, V. Morellas, and N. Papanikolopoulos, “Action Recognition Using Global Spatio-Temporal Features Derived from Sparse Representations”, *Computer Vision and Image Understanding*, 123:1-13, 2013

Workshop Papers

103. K. Kogashi and A. Cherian, MMHOI: Complex 3D Multi-Human-Object Interaction Understanding, PhyHumans Workshop at CVPR 2026 (**Oral**)
104. Y. Ni*, S. Wen, P. Koniusz, **A. Cherian**, “Noise Consistency Regularization for Subject-driven Image Synthesis”, SyntaGen workshop, CVPR 2025
105. A. Singh, M. Jones, M. Chatterjee, K-C. Peng, **A. Cherian**, E. Learned-Miller, “Improving Open-World Object Localization by Discovering Background”, Domain Generalization Workshop, CVPR, 2025
106. F. Mumcu, M. Jones, Y. Yilmaz, **A. Cherian**, “ComplexVAD: Detecting Interaction Anomalies in Video”, WACV Workshops, 2025
107. M. Chatterjee*, N. Ahuja, **A. Cherian**, “Quantifying Predictive Uncertainty for Stochastic Video Synthesis from Audio”, Sound and Sight CVPR Workshop, 2022
108. R. S. Cruz*, A. Cherian, B. Fernando, D. Campbell, and S. Gould, “Inferring Temporal Compositions of Actions Using Probabilistic Automata”, CVPR Workshops, 2020
109. Y. Wu, T. Marks, A. Cherian, S. Chen, C. Feng, G. Wang, and A. Sullivan, “Unsupervised Joint 3D Object Model Learning and 6D Pose Estimation for Depth-Based Instance Segmentation”, ICCV Workshops, 2019
110. H. Alamri, V. Cartillier, R. Gontijo Lopes, A. Das, J. Wang, I. Essa, D. Batra, D. Parikh, **A. Cherian**, T. K. Marks, C. Hori, “Audio Visual Scene-Aware Dialog (AVSD) Challenge at DSTC7”, CVPR Workshops, 2018
111. P. Stanitsas*, **A. Cherian**, V. Morellas, and N. Papanikolopoulos, “Clustering Positive Definite Matrices via Learning Information Divergences”, IEEE Workshop on Manifold Learning: From Euclid to Riemann, ICCV Workshops, Venice, 2017
112. J. Wang* and **A. Cherian**, “Discriminative Subspace Pooling for Action Recognition”, Workshop on Perceptual Organization in Computer Vision (POCV), European Conference on Computer Vision (ECCV), 2018

113. **A. Cherian** and N. Papanikolopoulos, “Large Scale Image Search via Sparse Coding”, Minnesota Supercomputing Institute (MSI) Poster Competition, 2012 (**Best Poster Award**)

Patents

1. M Jones, S Lohit, A Cherian, Z Carmichael, Systems and methods for interpretable classification of images using inherently explainable neural networks, 2025
2. M. Chatterjee, **A. Cherian**, Semantic Object Disambiguation And Robot Navigation Using A Hierarchical Scene Graph With Sensorial Constraints, 2025 (filed)
3. **A. Cherian**, Recursive Architecture for Graph-Based Auto-Encoder, 2025 (filed)
4. Y. Ni, **A. Cherian**, S. Wen, Method and System for Subject-Driven Image Synthesis with Noise Consistency Regularization, 2025 (filed)
5. M. Jones, **A. Cherian**, F. Mumcu, Y. Yilmaz, Video Anomaly Detection Based On Object Interactions, 2025 (filed)
6. **A. Cherian**, T. Marks, S. Jain, System and Method for Object Segmentation for Task Performance, 2024 (filed)
7. **A. Cherian**, Autoencoder with Non-Uniform Unrolling Recursion, 2024 (filed)
8. A Cherian, X Liu, S Paul, M Chatterjee, System and Method for Controlling a Robot, 2024
9. M Chatterjee, A Cherian, J Le Roux, Method and system for scene-aware audio-video representation, 2024
10. **A. Cherian**, “Anomaly detector for detecting anomaly using Complementary Classifiers”, US Patent, 2022
11. **A. Cherian**, “Spatio-Temporal Interaction Encoder for Video Anomaly Detection”, US Patent Application, 2022
12. **A. Cherian** and A. Shuchin, “Artificial Intelligence System for Classification of Data Based on Contrastive Learning”, US Patent, 2022
13. S. Geng, P. Gao, **A. Cherian**, C. Hori, and J. Le Roux, “Scene-Aware Video Dialog”, US Patent, 2021
14. C. Hori, **A. Cherian**, T. Marks, and T. Hori, “System and Method for a Dialogue Response Generation System”, US Patent, 2021
15. C. Hori, T. Hori, **A. Cherian**, T. Marks, J. Le Roux “Low-Latency Caption System”, US Patent, 2022
16. S. Paul and **A. Cherian**, System and Method for Controlling an Entity, 2022
17. **A. Cherian**, C. Hori, T. Marks, and J. Le Roux, “Scene-Aware Video Encoder System and Method”, 2022
18. M. Chatterjee, **A. Cherian**, and J. Le Roux, “Scene Aware Audio-Video Representation”, 2022
19. **A. Cherian**, G. Pais, S. Jain, T. Marks, and A. Sullivan, “A Generative Approach to Instance Segmentation in Depth Images”, 2022

20. **A. Cherian**, T. Marks, and A. Sullivan, “Discriminative 3D Shape Modeling for Few-Shot Instance Segmentation”, 2022
21. S. Medin, T. Marks, X. Liu, **A. Cherian**, B. Egger, and J. Tenenbaum, “System and Method for Manipulating Two-Dimensional (2D) Images of Three-Dimensional (3D) Objects”, 2022
22. C. Hori, **A. Cherian**, S. Chen, T. Marks, J. Le Roux, T. Hori, B. Hersham, A. Vetro, and A. Sullivan, “Method and System for Scene-Aware Interaction”, US Patent, 2021
23. C. Hori, M. Tsuchiya, S. Chen, **A. Cherian**, T. Hori, B. Hersham, T. Marks, J. Le Roux, A. Sullivan, and A. Vetro, “Scene-Aware Interaction”, US Patent, 2021
24. P. Stanitsas, **A. Cherian**, V. Morellas, N. Papanikolopoulos, and A. Truskinovsky, “Computer Vision for Cancerous Tissue Recognition”, US Patent, 2018

Invited Talks and Keynotes (short list)

- “On AI and Generalization” AI Business Interview, 2023
- “Are Deep Neural Networks SMARTer than Second Graders?”, Invited talk, at CogSci Lab, MIT, 2023
- “Learning a Constrained Optimizer: A Primal Approach”, AAAI CPML Workshop, 2023
- “Scene Representations for Multimodal Machine Intelligence”, Keynote, SANE Workshop 2022
- “InSeGAN: An Unsupervised Approach to Identical Instance Segmentation”, Invited talk, Visual Information Lab, University of Bristol, UK, 2021
- “Dynamic Graph Representation Learning for Audio-Visual Scene-Aware Dialog”, Invited talk at CogSci Lab, MIT, 2021
- “Sound2Sight: Audio-conditioned Visual Imagination”, Keynote, Multimodal Video Analysis Workshop, ECCV, 2020
- “Discriminative Video Representation Learning Using Adversarial Perturbations”, European Conference on Computer Vision (ECCV), Oral presentation, Munich, 2018
- “Fine-grained Action Representations for Human-Robot Interactions”, RoboVis, Melbourne, Sept, 2016
- “Tensor Sparse Representations for Fine-grained Action Recognition”, Mitsubishi Electric Research Labs, August, 2016
- “Tensor Representations for Fine-grained Action Recognition”, Invited Talk at CSAIL at MIT, July, 2016
- “Human Pose Estimation Using Recombination of Parts”, RoboVis, Adelaide, Australia, June 2015
- “Mixing Body-Part Sequences for Human Pose Estimation”, Microsoft Research-INRIA joint center, Paris, May 2014
- “Large Scale Image Search via Sparse Coding”, Minnesota Supercomputing Institute Research Exhibition, April 2012