European Journal of AI,
Computing & Informatics

Vol. 1 No. 2 2025

Review **Open Access**

Vision-Based AI Solutions for Human Life and Social Science: From Image Processing to Human Behavior Modeling

W. A. Jarvis [1,*]

[1] Department of Computer Science, Australian National University, Canberra, Australia
* Correspondence: W. A. Jarvis, Department of Computer Science, Australian National University, Canberra, Australia

Received: 29 June 2025
Revised: 08 July 2025
Accepted: 18 July 2025
Published: 26 July 2025

Abstract: Artificial Intelligence (AI) has become a transformative force in social science research, enabling the analysis of large-scale, heterogeneous data to uncover latent patterns and predict complex human behaviors. Among AI's core methodologies, computer vision has evolved far beyond its early role in data acquisition to now encompass sophisticated systems capable of interpreting, analyzing, and synthesizing visual information across a range of socially relevant contexts. By integrating advanced image processing, machine learning, and computer graphics, computer vision empowers interdisciplinary investigations in psychology, sociology, and economics, modeling phenomena such as decision-making, emotional expression, and social interaction at unprecedented scales and resolutions. This paper presents a comprehensive survey of the state of the art in computer vision applications within the social sciences, with particular emphasis on recent breakthroughs in algorithms and enabling technologies that facilitate automated visual understanding. Key topics include object detection, facial recognition, scene understanding, and predictive modeling, which collectively underpin impactful applications in healthcare, autonomous systems, surveillance, and digital media. To structure this rapidly expanding domain, we propose a conceptual framework that organizes the field into four foundational pillars: image processing, object recognition, adaptive machine learning, and computer graphics. Each pillar contributes critical capabilities such as feature extraction, quality enhancement, semantic interpretation, and photorealistic rendering — functions that are increasingly pivotal in addressing contemporary social challenges. By critically evaluating current methodologies, benchmarking performance across domains, and identifying emergent trends, this work not only synthesizes existing knowledge but also outlines promising directions for future research at the intersection of AI and social science. Finally, we highlight how these advancements are reshaping societal norms and enabling AI-driven solutions to pressing global issues, from autonomous navigation to public health and beyond.

Keywords: social science; computer vision; pattern recognition; artificial intelligence

1. Introduction

Artificial Intelligence (AI) has rapidly transformed human life, extending its influence beyond technological and industrial domains to encompass fundamental social structures and individual behaviors. Within this big landscape, computer vision emerges as a critical and important subfield, driving a wide range of applications that center on the knowledge, interpretation, and analysis of visual information within everyday human experiences. At its core, computer vision seeks to simulate the perceptual and cognitive capabilities of human vision, enabling machines to perceive, comprehend, and respond to visual data with increasing levels of complexity and autonomy. Given the rapid and extensive development of artificial intelligence applications [1-4], like face recognition by computer vision [5,6], multimedia system networks [7-16], generative adversarial networks [17-19], large language models [20-25], autonomous driving technologies [26-29], medical sciences [30,31], Spectroscopy[32,33], human health care[30,34-38], environmental &water protection[39-42], scientific research[34,40,43-45], radio-frequency identification[46,47] and design optimization[48] etc., Computer vision has expanded across diverse fields, evolving from basic raw data capture to advanced techniques for interpreting images and videos, fundamentally transforming how people interact with their surroundings and with one another. In practical applications in medical science, progress was made by computer vision in surgery of liver transplantation[49-54]. Recent achievements in social learning [55-62] provide even greater insights into real-world applications in transportation and traffic planning [63], which significantly impact human life. As one of the most outstanding application field, the very biggest progress has been achieved in computer vision[64-70] with powerful software[71,72] and hardware support[73-80] like millimeter Wave technology and Robotics[81]. This process requires techniques for feature extraction and event detection, which vary depending on the application domain and the nature of the data being analyzed.

In social science, computer vision provides new methods for analyzing human behavior, cultural traits, and social change by processing vast amounts of visual data from social media, urban spaces, and digital platforms. These provide a huge number of images used to train large models, which in turn rely on image processing and pattern recognition — both foundational components of computer vision. While image processing emphasizes techniques for enhancing image quality and extracting useful features, Pattern recognition focuses on identifying and categorizing objects or patterns within those images to reveal overall trends in human behavior. By combining these two areas, social science researchers and engineers can develop models capable of interpreting spatial data, ultimately leading to image-based understanding of human psychology. The field of computer vision draws significant inspiration from human behavior and visual capabilities, yet fully replicating human-like vision remains beyond current technological limits due to constraints in machine interpretation and processing. Although computer vision and image processing are closely related, they have distinct objectives. Computer vision aims to create models that extract relevant information and facilitate its interpretation, thus enabling image understanding. In contrast, image processing involves computational transformations — such as sharpening and contrast adjustment — that enhance the visual clarity of images. These fields occasionally intersect with Human-Computer Interaction (HCI) [82], an interdisciplinary area concerned with user-centered design, interface ergonomics, and effective human-technology engagement. Nevertheless, computer vision is specifically devoted to analyzing and understanding visual data from human daily life.

Both human and computer vision serve the same functional purpose: interpreting spatial data to understand objects, scenes, and actions in real time and real world [83]. Yet, despite ongoing advancements, computer vision still falls short of the human visual system's adaptability and contextual awareness. Developing algorithms that can capture the accuracy, flexibility, and complexity of human perception remains a significant challenge. Researchers grapple with various obstacles, including the fine-tuning of parameters, ensuring algorithmic robustness under diverse conditions, and maintaining reliable outcomes across multiple scenarios. The performance of computer vision systems also presents complexities and involves comprehensive assessments of accuracy, robustness, and adaptability. Such evaluations test how algorithms respond to different conditions and measure their ability to sustain performance over time. In light of these challenges, the field dedicates considerable effort to refining computer vision models. These applications range from industrial automation and remote sensing to robotics, human-computer interaction, and assistive technologies for individuals with visual impairments. Ultimately, advances in computer vision depend on improvements in computational power and algorithmic design. As technology continues to evolve, the scope of computer vision expands, providing novel ways to interpret and analyze visual information. However, the widespread use of computer vision also raises ethical concerns, particularly regarding privacy, surveillance, and algorithmic biases that can reinforce social inequalities. This survey underscores the vital role of computer vision in closing the gap between human visual capacity and machine-based analysis, driving innovation across numerous industries, and shaping human life and behavior.

2. Development and Impacts

The major development of computer vision leverages advanced algorithms and optical or inertial sensors to replicate certain capabilities of human vision [73,84-86], enabling the automatic extraction of valuable information from human behavior. In contrast to traditional techniques, which can be time-consuming and often necessitate complex laboratory analysis, computer vision can interpret visual data with greater speed and efficiency. This technology is often coupled with specialized lighting setups to optimize image capture and improve analytic accuracy. By acquiring digital images and refining their quality through preprocessing, computer vision systems can isolate target objects from backgrounds, measure key features, and then accurately interpret the resulting data.

The rapid expansion of social media platforms has led to an overwhelming influx of human-centric images, creating a significant demand for advanced image processing technologies. Recent progress in image processing has led to the development of exceptionally accurate digital recognition systems, revolutionizing how visual data is analyzed across a wide range of sectors [87]. These advancements streamline data collection and refine image interpretation, paving the way for new applications such as automated quality control, medical imaging, and autonomous navigation [88]. As a result, computer vision has evolved into a powerful technology capable of rapidly processing images and extracting crucial information, thereby offering a structured and efficient alternative to traditional visual analysis. Several examples highlight these innovations in image processing. One instance involves using precision metrology and digital image processing to measure the groove shapes of phonograph recordings without direct contact, which then facilitates the visual reconstruction of audio signals. Another application addresses the analysis of food and beverage imagery, employing feature selection, extraction, and classification methods to evaluate role portrayal and industry impact. Finally, convolutional Neural Networks (CNN) for object detection demonstrate how artificial neural networks (ANN) can effectively handle edge detection in various image processing tasks [89].

Digital image processing has evolved into a leading discipline within computer vision, largely due to progress in mathematics, linear algebra, statistics, scientific computing, and computational neuroscience. These fields furnish the theoretical underpinnings for developing and refining image processing techniques. Recent advances include depth map estimation methods that utilize Bayesian techniques [90] to recover 3D structures, yielding strong performance with training data but encountering challenges in natural image settings [91]. Similarly, image quality assessment for retargeting methods benefits from top-down approaches, which ensure consistent outcomes based on specific metrics. Research in this area also extends to psychophysical studies on human visual perception under varying lighting conditions, facilitated by dynamic displays and HDR technology, enabling detailed pixel differentiation across a broad luminance range. In practical applications, driver assistance systems employ contrast sensitivity techniques to assess driver visual perception in low-light conditions, thereby offering real-time guidance for speed adjustments in low-visibility scenarios. Image-based illumination enhancement represents another significant development: color pixel correction and decomposition surpass traditional histogram equalization by producing clearer images in low-light or high-contrast environments. Pattern recognition, a core component of computer vision, concentrates on object identification through image transformation techniques that enhance visual quality and enable precise interpretation. By extracting information from sensor-based imagery and informing subsequent decisions, it aspires to replicate aspects of human vision. The typical computer vision workflow includes image acquisition, preprocessing, feature extraction, segmentation or detection, high-level processing, and eventual decision-making.

Convolutional neural network–based object detection has gained significant prominence in various applications and research areas [92,93]. In particular, two-stage detectors such as the Region-based Convolutional Neural Network (R-CNN) family [94] are frequently highlighted. These multi-step processes form a coherent pipeline for handling visual data and accurately identifying objects. Within many computer vision frameworks, two primary methodologies are employed: 3D morphological analysis and pixel optimization. While 3D morphological analysis has become a mainstay for image processing and pattern recognition, Pixel optimization focuses on the detailed examination of pixel structures — including their morphology and inherent properties — to improve the understanding of vector-based image representations. Often, both methods are applied to extensive datasets capturing multiple layers of geometric composition. This approach demands efficient and precise algorithms that can effectively extract crucial quantitative details and interpret complex color clusters.

Combining 3D morphological analysis with artificial intelligence techniques [65] — such as fuzzy logic, artificial neural networks, and genetic algorithms — further enhances the performance of computer vision solutions, particularly in large-scale data scenarios. These AI-driven methods enable intricate tasks such as large-scale visual classification and real-time image segmentation that would otherwise require substantial human input. In practical applications, computer vision commonly employs two central processes for handling image data: segmentation and retrieval. Segmentation divides an image into separate regions of pixels that share specific characteristics (e.g., color, texture, or grayscale level), a critical step for accurate object detection and interpretation. It isolates each region for independent analysis, underpinning applications involving object recognition and classification. Retrieval then uses these segmented regions to locate visually similar images, forming the basis of content-based image retrieval systems and enriching the capabilities of image-based search engines. Collectively, segmentation and retrieval enable comprehensive data processing and categorization across a broad range of applications — from automated quality inspection to advanced search functionalities. Image segmentation frequently employs techniques based on intensity, color, and shape, as well as edge or border detection, all contributing to more accurate object recognition. Although segmentation accuracy is typically illustrated using small image samples in research, large-scale image databases often require specialized parameters for optimal classification. Advanced methods may incorporate gradient texture analysis, feature space exploration, and unsupervised clustering to precisely localize objects and delineate boundaries. Ultimately, segmentation aims to generate a resemblance map from established object detection models or hierarchical segmentation procedures, supporting a salience mapping framework that highlights key regions in an image. Vision-based models typically compute pixel-level salience values and map them to spatial coordinates within a hierarchical saliency map to highlight regions of interest. In some cases, researchers employ aggregation models that use established saliency techniques to assign salience scores to each pixel and segment, grouping them into prominent clusters.

Nevertheless, such aggregation models can overlook the interplay between neighboring pixels, which may reduce segmentation accuracy in densely populated or complex visual scenarios. For instance, while pixel-wise aggregation aids in establishing model parameters, it may neglect local dependencies critical for delineating object boundaries and transitions in detailed image structures. To address these gaps, current research in computer vision segmentation is exploring more sophisticated models that factor in the relational dynamics among pixels, thereby enhancing the reliability and depth of visual analysis across diverse applications. As computer vision continues to advance, these innovations in segmentation and pattern recognition serve as the foundation for more adaptive, context-aware systems, enabling them to navigate and interpret real-world environments with increasing precision.

To address these challenges, Conditional Random Fields (CRF) is proposed to aggregate calibration maps from multiple methods while incorporating values from neighboring pixels [95]. This CRF-based approach enhances parameter optimization during training by emphasizing the reliability of each pixel within the framework. Data extraction involves capturing objects through cameras, sensors, or satellite devices, either as single images or image sequences. The primary objective is to distinguish foreground objects from background elements. The process can produce three types of outputs: (a) original color retention, (b) grayscale conversion, and (c) transparency rendering. This technique can identify human emotions and predict behavior and habits in certain contexts.

The feature extraction process in computer vision includes multiple steps: (a) converting object instances to black and white, (b) resizing objects based on a scale factor, (c) applying transparency or color combinations to specific elements, and (d) scaling and repositioning objects — often resulting in appearances that differ from their original forms. Now, we can add one more step regarding human behavior identification or emotional change. On the other hand, pixel-level precision is critical in determining image sharpness, making optimization essential for tasks such as object detection, segmentation, and recognition. Boundary-based techniques employ edge detectors to delineate object boundaries by identifying rapid intensity changes along region edges. In color segmentation, this process is performed on each RGB channel, with the resulting edges combined to produce the final edge image. In contrast, local-based techniques group pixels based on uniformity criteria, as seen in region-growing and split-and-merge methods. Region-growing expands from core points into larger areas when pixels share similar features, such as color or grayscale values. Conversely, split-and-merge techniques initially subdivide an image into smaller regions that are subsequently merged according to specific criteria. However, these region-based methods face two main limitations: they heavily rely on initial global criteria that influence regional growth, and they depend on the initial segments and original pixel values, which can affect the accuracy of object detection. While object detection is widely applied to both recorded and real-time datasets, its accuracy can decrease when target objects differ significantly from predefined algorithmic patterns. To improve accuracy, supplementary algorithms are often deployed to detect finer features; for instance, face detection systems typically identify lower-resolution facial elements, such as eyes, eyebrows, and mouths, thereby enhancing overall precision. Peripheral features such as the ears and neck area, however, tend to receive less analytical attention.

A bitmap representation depicts an image as a collection of pixels, each assigned a specific color, and is commonly used for photographs and digital images. However, bitmap quality degrades when enlarged: scaling a bitmap image, say by a factor of four, results in enlarged pixels that lead to a blurred or pixelated appearance. Critical parameters when working with bitmap images include resolution, which refers to the total number of pixels, and color depth, indicating the range of colors each pixel can display. Bitmap images are generated using scanners, digital cameras, or video capture devices and are vulnerable to various types of noise. To mitigate this, standardized bitmap templates — easily processed by computers — are often used as benchmarks. In contrast, raw bitmap images may contain acquisition errors, resulting in unstructured pixel values that do not accurately represent the original scene. Noise can enter bitmap images through different mechanisms: for instance, scanning a film photograph may introduce noise from film grain, damage, or scanner interference, while images captured directly in digital format may be affected by CCD detectors or electronic data transmission issues.

The research regarding image processing is focused on developing machine learning and computational methods capable of recognizing patterns in increasingly diverse objects. Machine learning plays a crucial role in applications such as spam filtering, search engines, and computer vision. Various algorithms reduce noise by averaging pixel values within defined regions, thereby mitigating variations and producing clearer, more accurate images. As social media continues to evolve, image processing remains at the forefront of technological advancements, shaping how digital content is created, shared, and interpreted. This ongoing development opens new avenues for artificial intelligence (AI)-driven applications, enabling more efficient, secure, and user-friendly experiences across digital platforms.

3. Future Work

Understanding the fundamental impact of computer vision on social structures and human behavior is essential for both harnessing its vast potential and ensuring its responsible use while managing associated risks for scientific and societal advancement. Evolving from simple data capture to advanced systems capable of interpreting and analyzing visual information, computer vision has become a foundational discipline driving progress across diverse domains. It enables the conversion of raw images and videos into actionable insights, advancing applications in remote sensing, robotics, healthcare (e.g., COVID-19 diagnostics), and satellite communications. Beyond traditional pattern detection and object recognition, recent breakthroughs in machine learning have equipped computer vision systems with greater adaptability, real-time performance, and resilience in complex environments. These developments are paving the way for integration with emerging technologies such as edge computing, augmented reality, and mixed reality. As the field advances, future efforts will likely prioritize optimizing computational efficiency, improving robustness to environmental variability, and addressing the challenges posed by large-scale deployment. Importantly, the expanding influence of computer vision on employment, social interaction, and communication underscores the need for ethical frameworks that safeguard privacy, promote fairness, and balance technological progress with societal values. Continuous innovation in this field holds the promise of not only enhancing the accuracy and speed of visual data analysis but also reshaping how humans interact with technology, fostering data-driven solutions that advance human well-being and support the modeling and analysis of complex social behaviors [96].

4. Conclusion

In summary, this review has explored the transformative role of computer vision in social sciences, highlighting its capabilities in processing and interpreting visual data to model human behaviors and societal trends. By structuring the field into key pillars such as image processing, pattern recognition, and adaptive machine learning, this paper provides a clear framework for researchers and practitioners. Continued innovation in this domain promises to further bridge the gap between human perception and machine understanding, enabling responsible and impactful applications across diverse sectors.

References

1. Y. Wang, Y. Guo, R. Kumar, and M. Swaminathan, "Order reduction using Laguerre-FDTD with embedded neural network," in *Proc. IEEE/MTT-S Int. Microw. Symp. (IMS)*, 2024, pp. 473–476, doi: 10.1109/IMS40175.2024.10600239.
2. Z. An, G. Sun, Y. Liu, R. Li, M. Wu, M.-M. Cheng, E. Konukoglu, and S. Belongie, "Multimodality helps few-shot 3D point cloud semantic segmentation," *arXiv Prepr.* arXiv:2410.22489, 2024, doi: 10.48550/arXiv.2410.22489.
3. R. Li, J. Han, L. Melas-Kyriazi, C. Sun, Z. An, Z. Gui, S. Sun, P. Torr, and T. Jakab, "DreamBeast: Distilling 3D fantastical animals with part-aware knowledge transfer," *arXiv Prepr.* arXiv:2409.08271, 2024, doi: 10.48550/arXiv.2409.08271.
4. Y. Guo, X. Jia, X. Li, Y. Wang, R. Kumar, R. Sharma, and M. Swaminathan, "Extrapolation with range determination of 2D spectral transposed convolutional neural network for advanced packaging problems," *IEEE Trans. Compon. Packag. Manuf. Technol.*, vol. 13, no. 10, pp. 1533–1544, 2023, doi: 10.1109/TCPMT.2023.3317851.

5. R. Li, S. Sun, M. Elhoseiny, and P. Torr, “OxfordTVG-HIC: Can machine make humorous captions from images?,” in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, 2023, pp. 20293–20303, doi: 10.1109/ICCV51070.2023.01856.
6. X. Chen, K. He, W. Liu, X. Liu, Z.-J. Zha, and T. Mei, “CLaM: An open-source library for performance evaluation of text-driven human motion generation,” in *Proc. ACM Int. Conf. Multimedia (ACM MM)*, 2024, pp. 11194–11197, doi: 10.1145/3664647.3685523.
7. X. Chen, W. Liu, X. Liu, Y. Zhang, and T. Mei, “A cross-modality and progressive person search system,” in *Proc. ACM Int. Conf. Multimedia (ACM MM)*, 2020, pp. 4550–4552, doi: 10.1145/3394171.3414455.
8. X. Chen, X. Liu, K. Liu, W. Liu, and T. Mei, "A baseline framework for part-level action parsing and action recognition," *arXiv preprint arXiv:2110.03368*, 2021, doi: 10.48550/arXiv.2110.03368.
9. X. Chen, X. Liu, W. Liu, K. Liu, D. Wu, Y. Zhang, and T. Mei, "Part-level action parsing via a pose-guided coarse-to-fine framework," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, 2022, pp. 419–423, doi: 10.1109/ISCAS48785.2022.9937498.
10. W. Wang, Y. Sun, Z. Yang, Z. Hu, Z. Tan, and Y. Yang, "Replication in visual diffusion models: A survey and outlook," *arXiv preprint arXiv:2408.00001*, 2024, doi: 10.48550/arXiv.2408.00001.
11. X. Chen and X. Di, "Ridesharing user equilibrium with nodal matching cost and its implications for congestion tolling and platform pricing," *Transp. Res. Part C Emerg. Technol.*, vol. 129, p. 103233, 2021, doi: 10.1016/j.trc.2021.103233.
12. X. Chen and X. Di, "A network equilibrium model for integrated shared mobility services with ride-pooling," *Transp. Res. Part C Emerg. Technol.*, vol. 167, p. 104837, 2024, doi: 10.1016/j.trc.2024.104837.
13. C. Zhang, X. Chen, and X. Di, "Stochastic semi-gradient descent for learning mean field games with population-aware function approximation," *arXiv preprint arXiv:2408.08192*, 2024, doi: 10.48550/arXiv.2408.08192.
14. Q. Jin, X. Chen, W. Liu, T. Mei, and Y. Zhang, "T-SVG: Text-driven stereoscopic video generation," *arXiv preprint arXiv:2412.09323*, 2024, doi: 10.48550/arXiv.2412.09323.
15. M. Yin, "Multimedia authentication for copyright protection," *IOP Conf. Ser. Earth Environ. Sci.*, IOP Publishing, 2017, p. 012160, doi: 10.1088/1755-1315/69/1/012160.
16. Y. Deng, S. Shao, A. Mittal, R. Twumasi-Boakye, J. Fishelson, A. Gupta, and N. B. Shroff, "Incentive design and profit sharing in multi-modal transportation networks," *Transp. Res. Part B Methodol.*, vol. 163, pp. 1–21, 2022, doi: 10.1016/j.trb.2022.06.011.
17. Y. Hu, M. Yin, M. Mezzavilla, H. Guo, and S. Rangan, "Channel modeling for FR3 upper mid-band via generative adversarial networks," in *Proc. 2024 IEEE 25th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Sep. 2024, pp. 776–780, doi: 10.1109/SPAWC60668.2024.10693976.
18. Y. Hu, M. Yin, S. Rangan, and M. Mezzavilla, "Parametrization and estimation of high-rank line-of-sight MIMO channels with reflected paths," *IEEE Trans. Wireless Commun.*, 2023, doi: 10.1109/TWC.2023.3311735.
19. Y. Hu, M. Yin, S. Rangan, and M. Mezzavilla, "Parametrization of high-rank line-of-sight MIMO channels with reflected paths," in *Proc. IEEE 23rd Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, 2022, pp. 1–5, doi: 10.1109/SPAWC51304.2022.9833962.
20. M. Qu, X. Chen, W. Liu, A. Li, and Y. Zhao, "ChatVTG: Video temporal grounding via chat with video dialogue large language models," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2024, pp. 1847–1856, doi: 10.1109/CVPRW63382.2024.00191.
21. L. Yang, Z. Zhang, J. Han, B. Zeng, R. Li, P. Torr, and W. Zhang, "Semantic score distillation sampling for compositional text-to-3D generation," *arXiv preprint arXiv:2410.09009*, 2024, doi: 10.48550/arXiv.2410.09009.
22. Z. Gui, S. Sun, R. Li, J. Yuan, Z. An, K. Roth, A. Prabhu, and P. Torr, "kNN-CLIP: Retrieval enables training-free segmentation on continually expanding large vocabularies," *arXiv preprint arXiv:2404.09447*, 2024, doi: 10.48550/arXiv.2404.09447.
23. B. Wang, H. Duan, Y. Feng, X. Chen, Y. Fu, Z. Mo, and X. Di, "Can LLMs understand social norms in autonomous driving games?," *arXiv preprint arXiv:2408.12680*, 2024, doi: 10.1109/IAVVC63304.2024.10786452.
24. H. Liu, X. Chen, X. Liu, X. Gu, and W. Liu, "AnimateAnywhere: Context-controllable human video generation with ID-consistent one-shot learning," in *Proc. 5th Int. Workshop Human-centric Multimedia Anal.*, 2024, pp. 41–43, doi: 10.1145/3688865.3689477.
25. Y. Mohamed, R. Li, I. S. Ahmad, K. Haydarov, P. Torr, K. W. Church, and M. Elhoseiny, "No culture left behind: ArtELingo-28, a benchmark of WikiArt with captions in 28 languages," *arXiv preprint arXiv:2411.03769*, 2024, doi: 10.48550/arXiv.2411.03769.
26. M. Yin, T. Li, H. Lei, Y. Hu, S. Rangan, and Q. Zhu, "Zero-shot wireless indoor navigation through physics-informed reinforcement learning," in *Proc. 2024 IEEE Int. Conf. Robot. Autom. (ICRA)*, 2024, pp. 5111–5118, doi: 10.1109/ICRA57147.2024.10611229.
27. K. Huang, X. Chen, X. Di, and Q. Du, "Dynamic driving and routing games for autonomous vehicles on networks: A mean field game approach," *Transp. Res. Part C Emerg. Technol.*, vol. 128, p. 103189, 2021, doi: 10.1016/j.trc.2021.103189.
28. Y. Deng, H. Chen, S. Shao, J. Tang, J. Pi, and A. Gupta, "Multi-objective vehicle rebalancing for ridehailing system using a reinforcement learning approach," *J. Manage. Sci. Eng.*, vol. 7, no. 2, pp. 346–364, 2022, doi: 10.1016/j.jmse.2021.12.004.
29. Y. Deng, A. Gupta, and N. B. Shroff, "Fleet sizing and charger allocation in electric vehicle sharing systems," *IFAC J. Syst. Control*, vol. 22, p. 100210, 2022, doi: 10.1016/j.ifacsc.2022.100210.
30. J. Huo, H. Li, J. Roveda, S. F. Quan, and A. Li, "A multi-task deep learning algorithm for sleep stage scoring and sleep arousal detection," *Authorea Preprints*, 2023, doi: 10.36227/techrxiv.24078252.v1.

31. X. Chen and X. Di, "How the COVID-19 pandemic influences human mobility? Similarity analysis leveraging social media data," in *Proc. 2022 IEEE 25th Int. Conf. Intell. Transp. Syst. (ITSC)*, 2022, pp. 2955–2960, doi: 10.1109/ITSC55140.2022.9922060.
32. H. Guo, A. B. Tikhomirov, A. Mitchell, I. P. J. Alwayn, H. Zeng, and K. C. Hewitt, "Real-time assessment of liver fat content using a filter-based Raman system operating under ambient light through lock-in amplification," *Biomed. Opt. Express*, vol. 13, no. 10, pp. 5231–5245, 2022, doi: 10.1364/BOE.467849.
33. H. Guo, J. Huo, and Q. Jin et al., "Liver discard rate due to conservative estimations of steatosis: An inference-based approach," *medRxiv Prepr*., 2023, doi: 10.1101/2023.12.04.23299406.
34. C. Ding, T. Yao, C. Wu, and J. Ni, "Deep learning for personalized electrocardiogram diagnosis: A review," *arXiv preprint arXiv:2409.07975*, 2024, doi: 10.48550/arXiv.2409.07975.
35. J. Huo, S. F. Quan, J. Roveda, and A. Li, "BASH-GN: A new machine learning–derived questionnaire for screening obstructive sleep apnea," *Sleep Breath.*, vol. 27, no. 2, pp. 449–457, 2023, doi: 10.1007/s11325-022-02629-8.
36. J. Yang, "Research on the propagation model of COVID-19 based on virus dynamics," in *Proc. 2nd Int. Conf. Biol. Eng. Med. Sci. (ICBioMed)*, SPIE, 2023, pp. 962–967, doi: 10.1117/12.2669681.
37. J. Huo, S. F. Quan, J. Roveda, and A. Li, "Coupling analysis of heart rate variability and cortical arousal using a deep learning algorithm," *PLoS One*, vol. 18, no. 4, p. e0284167, 2023, doi: 10.1371/journal.pone.0284167.
38. M. Abdelkhalek, J. Qiu, M. Hernandez, A. Bozkurt, and E. Lobaton, "Investigating the relationship between cough detection and sampling frequency for wearable devices," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, 2021, pp. 7103–7107, doi: 10.1109/EMBC46164.2021.9630082.
39. J. Yang, "Predicting water quality through daily concentration of dissolved oxygen using improved artificial intelligence," *Sci. Rep.*, vol. 13, no. 1, p. 20370, 2023, doi: 10.1038/s41598-023-47060-5.
40. Y. Deng, X. Zhou, B. Kim, A. Tewari, A. Gupta, and N. B. Shroff, "Weighted Gaussian process bandits for non-stationary environments," *arXiv preprint arXiv:2107*, 2021, doi: 10.48550/arXiv.2107.02371.
41. S. Jiao and J. J. McCarthy, "A synergistic approach to atmospheric water scavenging," *ACS Appl. Mater. Interfaces*, vol. 15, no. 5, pp. 7353–7358, 2023, doi: 10.1021/acsami.2c18920.
42. S. Jiao and J. McCarthy, "Novel composites for atmospheric water absorption," in *Proc. 2020 Virtual AIChE Annu. Meeting*, AIChE, 2020.
43. J. Huo, Y. Wang, N. Wang, W. Gao, J. Zhou, and Y. Cao, "Data-driven design and optimization of ultra-tunable acoustic metamaterials," *Smart Mater. Struct.*, vol. 32, no. 5, p. 05LT01, 2023, doi: 10.1088/1361-665X/acc36c.
44. Y. Deng, X. Zhou, A. Ghosh, A. Gupta, and N. B. Shroff, "Interference constrained beam alignment for time-varying channels via kernelized bandits," in *Proc. 20th Int. Symp. Modeling Optim. Mobile, Ad Hoc, Wireless Netw. (WiOpt)*, IEEE, 2022, pp. 25–32, doi: 10.23919/WiOpt56218.2022.9930591.
45. M. Mieles, A. D. Walter, S. Wu, Y. Zheng, G. R. Schwenk, M. W. Barsoum, and H. F. Ji, "Hydronium-crosslinked inorganic hydrogel comprised of 1D lepidocrocite titanate nanofilaments," *Adv. Mater.*, vol. 36, no. 50, p. 2409897, 2024, doi: 10.1002/adma.202409897.
46. C. Qi, F. Amato, Y. Guo, Y. Zhang, and G. D. Durgin, "A backscatter channel sounder using tunneling RFID tags," in *Proc. 2021 IEEE Int. Conf. RFID (RFID)*, IEEE, 2021, pp. 1–7, doi: 10.1109/RFID52461.2021.9444368.
47. Y. Wang, Y. Guo, R. Kumar, R. Sharma, and M. Swaminathan, "Edge-based material cell meshing for improved accuracy of Laguerre-FDTD method," in *Proc. 2023 IEEE Int. Symp. Antennas Propag. USNC-URSI Radio Sci. Meeting (USNC-URSI)*, IEEE, 2023, pp. 1429–1430, doi: 10.1109/USNC-URSI52151.2023.10238031.
48. Y. Guo, O. W. Bhatti, and M. Swaminathan, "Training set optimization with uncertainty quantification for machine learning models of electromagnetic structures," in *Proc. 2022 IEEE Electr. Des. Adv. Packag. Syst. (EDAPS)*, IEEE, 2022, pp. 1–3, doi: 10.1109/EDAPS56906.2022.9994897.
49. H. Guo, A. E. Stueck, J. B. Doppenberg, Y. S. Chae, A. B. Tikhomirov, H. Zeng, M. A. Engelse, B. L. Gala-Lopez, A. Mahadevan-Jansen, and I. P. Alwayn, "Evaluation of minimum-to-severe global and macrovesicular steatosis in human liver specimens: a portable ambient light-compatible spectroscopic probe," *medRxiv*, 2023, doi: 10.1002/jbio.202400292.
50. H. Guo, V. S. Zions, B. A. Law, and K. C. Hewitt, "Potential of Raman-reflectance combination in quantifying liver steatosis and fat droplet size: evidence from Monte Carlo simulations and phantom studies," *J. Biophotonics*, 2024, Art. no. e202400156, doi: 10.1002/jbio.202400156.
51. H. Guo, A. E. Stueck, J. B. Doppenberg, Y. S. Chae, A. B. Tikhomirov, H. Zeng, B. L. Gala-Lopez, A. Mahadevan-Jansen, M. A. Engelse, and I. P. Alwayn, "Assessment of liver steatosis using an ambient light-compatible Raman system: enhancing specificity with supplementary reflectance information," in *Proc. Biomed. Vibrational Spectrosc. 2024: Adv. Res. Ind.*, SPIE, 2024, p. PC128390B, doi: 10.1117/12.3009086.
52. H. Guo, A. E. Stueck, A. B. Tikhomirov, H. Zeng, I. P. Alwayn, B. L. Gala-Lopez, A. Mahadevan-Jansen, A. K. Locke, and K. C. Hewitt, "Evaluation of steatosis in human liver specimens using an ambient light-compatible Raman spectroscopy approach," in *Proc. Bio-Optics: Des. Appl.*, Optica Publishing Group, 2023, p. JTu4B.26, doi: 10.1364/BODA.2023.JTu4B.26.
53. K. Jo, J. Choi, K. Kim, S. Kim, D. L. Nguyen, X. T. Vo *et al.*, "Artificial Behavior Intelligence: Technology, Challenges, and Future Directions," *arXiv preprint* arXiv:2505.03315, 2025, doi: 10.48550/arXiv.2505.03315.

54. E. Murphy-Chutorian, A. Doshi, and M. M. Trivedi, "Head pose estimation for driver assistance systems: A robust algorithm and experimental evaluation," in Proc. IEEE Intell. Transp. Syst. Conf. (ITSC), Sept. 2007, pp. 709–714, doi: 10.1109/ITSC.2007.4357803.
55. Z. Shou, X. Chen, Y. Fu, and X. Di, "Multi-agent reinforcement learning for Markov routing games: a new modeling paradigm for dynamic traffic assignment," *Transp. Res. Part C Emerg. Technol.*, vol. 137, p. 103560, 2022, doi: 10.1016/j.trc.2022.103560.
56. S. Liu, Y. Wang, X. Chen, Y. Fu, and X. Di, "SMART-eFlo: An integrated SUMO-gym framework for multi-agent reinforcement learning in electric fleet management problem," in *Proc. IEEE Int. Conf. Intell. Transp. Syst. (ITSC)*, IEEE, 2022, pp. 3026–3031, doi: 10.1109/ITSC55140.2022.9922047.
57. X. Chen, S. Liu, and X. Di, "A hybrid framework of reinforcement learning and physics-informed deep learning for spatiotemporal mean field games," in *Proc. 20th Int. Conf. Autonomous Agents Multiagent Syst.*, ACM Digital Library, 2023.
58. F. Zhou, C. Zhang, X. Chen, and X. Di, "Graphon mean field games with a representative player: Analysis and learning algorithm," *arXiv* preprint arXiv:2405.08005, 2024, doi: 10.48550/arXiv.2405.08005.
59. X. Chen, S. Liu, and X. Di, "Learning dual mean field games on graphs," in *Proc. Eur. Conf. Artif. Intell. (ECAI)*, 2023, pp. 421–428, doi: 10.3233/FAIA230299.
60. S. Liu, X. Chen, and X. Di, "Scalable learning for spatiotemporal mean field games using physics-informed neural operator," *Mathematics*, vol. 12, no. 6, p. 803, 2024, doi: 10.3390/math12060803.
61. X. Chen, Z. Li, and X. Di, "Social learning in Markov games: Empowering autonomous driving," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, IEEE, 2022, pp. 478–483, doi: 10.1109/IV51971.2022.9827289.
62. X. Chen, X. Di, and Z. Li, "Social learning for sequential driving dilemmas," *Games*, vol. 14, no. 3, p. 41, 2023, doi: 10.3390/g14030041.
63. X. Chen and X. Di, "Legal framework for rear-end crashes in mixed-traffic platooning: A matrix game approach," *Future Transp.*, vol. 3, no. 2, pp. 417–428, 2023, doi: 10.3390/futuretransp3020025.
64. Z. Hu, Y. Sun, and Y. Yang, "Switch to generalize: Domain-switch learning for cross-domain few-shot classification," in *Proc. Int. Conf. Learn. Representations (ICLR)*, 2022.
65. Z. Hu, Y. Sun, and Y. Yang, "Suppressing the heterogeneity: A strong feature extractor for few-shot segmentation," in *Proc. 11th Int. Conf. Learn. Representations (ICLR)*, 2023.
66. X. Chen, X. Liu, W. Liu, X.-P. Zhang, Y. Zhang, and T. Mei, "Explainable person re-identification with attribute-guided metric distillation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, 2021, pp. 11813–11822, doi: 10.1109/ICCV48922.2021.01160.
67. Z. Hu, Y. Sun, Y. Yang, and J. Zhou, "Divide-and-regroup clustering for domain adaptive person re-identification," in *Proc. AAAI Conf. Artif. Intell.*, 2022, pp. 980–988 , doi: 10.1609/aaai.v36i1.19981.
68. X. Chen, W. Liu, X. Liu, Y. Zhang, J. Han, and T. Mei, "MAPLE: Masked pseudo-labeling autoencoder for semi-supervised point cloud action recognition," in *Proc. 30th ACM Int. Conf. Multimedia*, 2022, pp. 708–718, , doi: 10.1145/3503161.3547892.
69. Z. Hu, Y. Sun, J. Wang, and Y. Yang, "DAC-DETR: Divide the attention layers and conquer," *Adv. Neural Inf. Process. Syst.*, vol. 36, pp. 75189–75200, 2023
70. X. Chen, W. Liu, Q. Bao, X. Liu, Q. Yang, R. Dai, and T. Mei, "Motion capture from inertial and vision sensors," *arXiv preprint* arXiv:2407.16341, 2024, doi: 10.48550/arXiv.2407.16341.
71. M. Yin, "Data security and privacy preservation in big data age," in *Proc. 2nd Int. Conf. Mechatronics Eng. Inf. Technol. (ICMEIT)*, Atlantis Press, 2017, pp. 387–391, doi: 10.2991/icmeit-17.2017.76.
72. J. Qiu and A. Aysu, "SHIFT SNARE: Uncovering Secret Keys in FALCON via Single-Trace Analysis," arXiv preprint arXiv:2504.00320, 2025, doi: 10.48550/arXiv.2504.00320.
73. M. Yin, A. K. Veldanda, A. Trivedi, J. Zhang, K. Pfeiffer, Y. Hu, S. Garg, E. Erkip, L. Righetti, and S. Rangan, "Millimeter wave wireless assisted robot navigation with link state classification," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 493–507, 2022, doi: 10.1109/OJCOMS.2022.3155572.
74. S. Yang, “The Impact of Continuous Integration and Continuous Delivery on Software Development Efficiency”, *J. Comput. Signal Syst. Res.*, vol. 2, no. 3, pp. 59–68, Apr. 2025, doi: 10.71222/pzvfqm21.
75. V. Semkin, M. Yin, Y. Hu, M. Mezzavilla, and S. Rangan, "Drone detection and classification based on radar cross section signatures," in *Proc. Int. Symp. Antennas Propag. (ISAP)*, IEEE, 2021, pp. 223–224, doi: 10.23919/ISAP47053.2021.9391260.
76. M. Yin, *Millimeter Wave Wireless Assisted Indoor Robot Navigation*, Doctoral dissertation, New York University Tandon School of Engineering, 2024.
77. K. Pfeiffer, Y. Jia, M. Yin, A. K. Veldanda, Y. Hu, A. Trivedi , et al., "Path planning under uncertainty to localize mmWave sources," arXiv preprint arXiv:2303.03739, 2023, doi:10.48550/arXiv.2303.03739.
78. Y. Hu, M. Yin, W. Xia, S. Rangan, and M. Mezzavilla, “Multi-frequency channel modeling for millimeter wave and THz wireless communication via generative adversarial networks,” in *Proc. 56th Asilomar Conf. Signals, Syst., Comput.*, IEEE, 2022, pp. 670–676, doi: 10.1109/IEEECONF56349.2022.10052063.
79. M. G. Aram, H. Guo, M. Yin, and T. Svensson, "Site-Specific Outdoor Propagation Assessment and Ray-Tracing Analysis for Wireless Digital Twins," in *Proc. 19th Eur. Conf. Antennas Propag. (EuCAP)*, Mar. 2025, pp. 1–5, IEEE, doi: 10.23919/EuCAP63536.2025.10999688.

80. F. Gao, "The Role of Data Analytics in Enhancing Digital Platform User Engagement and Retention", *J. Media Journal. Commun. Stud.*, vol. 1, no. 1, pp. 10–17, Apr. 2025, doi: 10.71222/z27xzp64.
81. S. Cao and J. Xiao, "Human-Robot Complementary Collaboration for Flexible and Precision Assembly," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2024, pp. 12971–12977, doi: 10.1109/ICRA57147.2024.10610825.
82. T. Kosch, J. Karolus, J. Zagermann, H. Reiterer, A. Schmidt, and P. W. Woźniak, "A survey on measuring cognitive workload in human-computer interaction," *ACM Comput. Surv.*, vol. 55, no. 13s, pp. 1–39, 2023 , doi: 10.1145/3582272
83. A. S. Abhigyan, M. Yin, and X. T. Tran, "Real-time unmanned aerial vehicle connectivity," U.S. Patent Appl. 17/502,568, 2023.
84. Z. Yuan, F. Lang, T. Xu, and X. Yang, "Sr-lio: Lidar-inertial odometry with sweep reconstruction," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2024, pp. 7862–7869, doi: 10.1109/IROS58592.2024.10802314.
85. Z. Yuan, J. Deng, R. Ming, F. Lang, and X. Yang, "SR-LIVO: LiDAR-inertial-visual odometry and mapping with sweep reconstruction," *IEEE Robot. Autom. Lett.*, 2024, doi: 10.1109/LRA.2024.3389415.
86. Z. Yuan, F. Lang, J. Deng, H. Luo, and X. Yang, "Voxel-svio: Stereo visual-inertial odometry based on voxel map," *IEEE Robot. Autom. Lett.*, 2025, doi: 10.1109/LRA.2025.3568307.
87. L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, and M. Pietikäinen, "Deep learning for generic object detection: A survey," *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 261–318, 2020, doi: 10.1007/s11263-019-01247-4.
88. E. Imani, G. Zhang, R. Li, J. Luo, P. Poupart, P. H. Torr, and Y. Pan, "Label Alignment Regularization for Distribution Shift," *arXiv preprint arXiv:2211.14960*, 2022 , doi: 10.48550/arXiv.2211.14960.
89. Y. Guo, X. Li, and M. Swaminathan, "2D spectral transposed convolutional neural network for S-parameter predictions," in *2022 IEEE 31st Conference on Electrical Performance of Electronic Packaging and Systems (EPEPS)*, 2022, pp. 1–3, doi: 10.1109/EPEPS53828.2022.9947109.
90. M. Swaminathan, O. W. Bhatti, Y. Guo, E. Huang, and O. Akinwande, "Bayesian learning for uncertainty quantification, optimization, and inverse design," *IEEE Trans. Microwave Theory Tech.*, vol. 70, no. 11, pp. 4620–4634, 2022, doi: 10.1109/TMTT.2022.3206455.
91. Y. Guo, X. Li, Y. Wang, R. Kumar, and M. Swaminathan, "Batch Training of Gaussian Process for Up-sampling Problems in S-Parameter Predictions," in *2023 IEEE 32nd Conference on Electrical Performance of Electronic Packaging and Systems (EPEPS)*, 2023, pp. 1–3, doi: 10.1109/EPEPS58208.2023.10314896.
92. Y. Fu, A. Jain, X. Chen, Z. Mo, and X. Di, "DriveGenVLM: Real-world Video Generation for Vision Language Model Based Autonomous Driving," in *2024 IEEE International Automated Vehicle Validation Conference (IAVVC)*, 2024, pp. 1-6, doi: 10.1109/IAVVC63304.2024.10786438.
93. X. Chen, F. Yongjie, S. Liu, and X. Di, "Physics-informed neural operator for coupled forward-backward partial differential equations," in *1st Workshop on the Synergy of Scientific and Machine Learning Modeling @ ICML 2023*, 2023.
94. S. Sun, R. Li, P. Torr, X. Gu, and S. Li, "Clip as rnn: Segment countless visual concepts without training endeavor," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition*, 2024, pp. 13171–13182, doi: 10.1109/CVPR52733.2024.01251.
95. M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (VOC) challenge," *Int. J. Comput. Vis.*, vol. 88, pp. 303–338, 2010, doi: 10.1007/S11263-009-0275-4.
96. B. Song, X. Wang, P. Sun, and A. Boukerche, "Robust COVID-19 vaccination control in a multi-city dynamic transmission network: A novel reinforcement learning-based approach," *J. Network Comput. Appl.*, vol. 219, 2023, Art. no. 103715, doi: 10.1016/j.jnca.2023.103715.