[['Distractor suppression Siamese network with task-aware attention for visual tracking',
  'https://www.sciencedirect.com/science/article/pii/S1077314225003303',
  '',
  'Zhigang Liu, Fuyuan Xing, Hao Huang, Kexin Wang, Yuxuan Shao',
  '1',
  '2026',
  '1'],
 ['Transformer tracking with high-low frequency attention',
  'https://www.sciencedirect.com/science/article/pii/S1077314225002863',
  '',
  'Zhi Chen, Zhen Yu',
  '1',
  '2026',
  '1'],
 ['PlanarTrack: A high-quality and challenging benchmark for large-scale planar object tracking',
  'https://www.sciencedirect.com/science/article/pii/S1077314225002760',
  '',
  'Yifan Jiao, Xinran Liu, Xiaoqiong Liu, Xiaohui Yuan, Heng Fan, Libo Zhang',
  '1',
  '2025',
  '12'],
 ['Curvi-Tracker: Curvilinear structure segmentation refinement by iterative tracking',
  'https://www.sciencedirect.com/science/article/pii/S0031320325014608',
  '',
  'Zhan Heng, Maurice Pagnucco, Erik Meijering, Yang Song',
  '2',
  '2026',
  '5'],
 ['Hybrid-stage association with dynamicity adaptation and enhanced cues for multi-object tracking and segmentation',
  'https://www.sciencedirect.com/science/article/pii/S0031320325014669',
  '',
  'Longtao Chen, Guoxing Liao, Yifan Shi, Jing Lou, Fenglei Xu, Huanqiang Zeng',
  '2',
  '2026',
  '5'],
 ['Spherical Vision Transformers for Audio-Visual Saliency Prediction in 360$^{\\circ }$° Videos',
  'http://ieeexplore.ieee.org/document/11144923',
  'Omnidirectional videos (ODVs) are redefining viewer experiences in virtual reality (VR) by offering an unprecedented full field-of-view (FOV). This study extends the domain of saliency prediction to 360$^\\circ$° environments, addressing the complexities of spherical distortion and the integration of spatial audio. Contextually, ODVs have transformed user experience by adding a spatial audio dimension that aligns sound direction with the viewer’s perspective in spherical scenes. Motivated by the lack of comprehensive datasets for 360$^\\circ$° audio-visual saliency prediction, our study curates YT360-EyeTracking, a new dataset of 81 ODVs, each observed under varying audio-visual conditions. Our goal is to explore how to utilize audio-visual cues to effectively predict visual saliency in 360$^\\circ$° videos. Towards this aim, we propose two novel saliency prediction models: SalViT360, a vision-transformer-based framework for ODVs equipped with spherical geometry-aware spatio-temporal attention layers, and SalViT360-AV, which further incorporates transformer adapters conditioned on audio input. Our results on a number of benchmark datasets, including our YT360-EyeTracking, demonstrate that SalViT360 and SalViT360-AV significantly outperform existing methods in predicting viewer attention in 360$^\\circ$° scenes. Interpreting these results, we suggest that integrating spatial audio cues in the model architecture is crucial for accurate saliency prediction in omnidirectional videos.',
  'Mert Cokelek, Halit Ozsoy, Nevrez Imamoglu, Cagri Ozcinar, Inci Ayhan, Erkut Erdem, Aykut Erdem',

```
  '3',
  '2026',
  '12'],
 ['PMGT-VR: A Decentralized Proximal-Gradient Algorithmic Framework With
Variance Reduction',
  'http://ieeexplore.ieee.org/document/11152599',
  'This article considers the decentralized composite optimization problem. We
propose a novel decentralized variance-reduction proximal-gradient algorithmic
framework, called PMGT-VR, which combines several techniques, including multi-
consensus, gradient tracking, and variance reduction. The proposed framework
imitates centralized algorithms and algorithms under this framework achieve
convergence rates similar to that of their centralized counterparts. We also
describe and analyze two representative algorithms, PMGT-SAGA and PMGT-LSVRG,
and compare them to existing state-of-the-art proximal algorithms. To the best
of our knowledge, PMGT-VR is the first linearly convergent decentralized
stochastic algorithm that can solve decentralized composite optimization
problems. Numerical experiments are provided to demonstrate the effectiveness of
the proposed algorithms.',
  'Haishan Ye, Wei Xiong, Tong Zhang',
  '3',
  '2026',
  '12'],
 ['ADA-Track++: End-to-End Multi-Camera 3D Multi-Object Tracking With
Alternating Detection and Association',
  'http://ieeexplore.ieee.org/document/11175504',
  'Many query-based approaches for 3D Multi-Object Tracking (MOT) adopt the
tracking-by-attention paradigm, utilizing track queries for identity-consistent
detection and object queries for identity-agnostic track spawning. Tracking-by-
attention, however, entangles detection and tracking queries in one embedding
for both the detection and tracking task, which is sub-optimal. Other approaches
resemble the tracking-by-detection paradigm and detect objects using decoupled
track and detection queries followed by a subsequent association. These methods,
however, do not leverage synergies between the detection and association task.
Combining the strengths of both paradigms, we introduce ADA-Track++, a novel
end-to-end framework for 3D MOT from multi-view cameras. We introduce a
learnable data association module based on edge-augmented cross-attention,
leveraging appearance and geometric features. We also propose an auxiliary token
in this attention-based association module, which helps mitigate
disproportionately high attention to incorrect association targets caused by
attention normalization. Furthermore, we integrate this association module into
the decoder layer of a DETR-based 3D detector, enabling simultaneous DETR-like
query-to-image cross-attention for detection and query-to-query cross-attention
for data association. By stacking these decoder layers, queries are refined for
the detection and association task alternately, effectively harnessing the task
dependencies. We evaluate our method on the nuScenes dataset and demonstrate the
advantage of our approach compared to the two previous paradigms.',
  'Shuxiao Ding, Lukas Schneider, Marius Cordts, Juergen Gall',
  '3',
  '2026',
  '12'],
 ['SNNTracker: Online High-Speed Multi-Object Tracking With Spike Camera',
  'http://ieeexplore.ieee.org/document/11165142',
  'Multi-object tracking (MOT) is crucial for applications such as autonomous
driving and robotics, yet traditional image-based methods struggle in high-speed
scenarios due to motion blur and temporal gaps caused by low frame rates. Spike
cameras, with their ability to continuously record spatiotemporal signals,
overcome these limitations. However, existing spike-based methods often rely on
intermediate image reconstruction or discrete clustering, limiting real-time
performance and temporal continuity. To address this, we propose SNNTracker, the
first fully spiking neural network (SNN)-based MOT algorithm tailored for spike
cameras. SNNTracker integrates a dynamic neural field (DNF)-based attention
mechanism for target detection and a winner-take-all (WTA)-based tracking module
with online spike-timing-dependent plasticity (STDP) for adaptive learning of
object trajectories. By directly processing spike streams without
```

reconstruction, SNNTracker reduces latency, computational overhead, and dependency on image quality, making it ideal for ultra-high-speed environments. It maintains robust, continuous tracking even under occlusions, severe lighting variations, or temporary object disappearance, by leveraging SNN-estimated motion predictions and long-term online clustering. We construct three types of spike-camera MOT datasets covering dense and sparse annotations across diverse real-world scenarios, including camera ego-motion, deformable and ultra-fast motion (up to 2600 RPM), occlusion, indoor/outdoor lighting changes, and low-visibility tracking. Extensive experiments demonstrate that SNNTracker consistently outperforms state-of-the-art MOT methods—both ANN- and SNN-based—achieving MOTA scores above 96% and up to 100% in many sequences. Our results highlight the advantages of spike-driven SNNs for low-latency, high-speed, and label-free multi-object tracking, advancing neuromorphic vision for real-time perception.',
  'Yajing Zheng, Chengen Li, Jiyuan Zhang, Zhaofei Yu, Tiejun Huang',
  '3',
  '2026',
  '12'],
 ['LVOS: A Benchmark for Large-Scale Long-Term Video Object Segmentation',
  'http://ieeexplore.ieee.org/document/11168273',
  'Video object segmentation (VOS) aims to distinguish and track target objects in a video. Despite the excellent performance achieved by off-the-shelf VOS models, part of the existing VOS benchmarks mainly focuses on short-term videos, where objects remain visible most of the time. However, these benchmarks may not fully capture challenges encountered in practical applications, and the absence of long-term datasets restricts further investigation of VOS in realistic scenarios. Thus, we propose a novel benchmark named LVOS, comprising 720 videos with 296,401 frames and 407,945 high-quality annotations. Videos in LVOS last 1.14 minutes on average. Each video includes various attributes, especially challenges encountered in the wild, such as long-term reappearing and cross-temporal similar objects. Compared to previous benchmarks, our LVOS better reflects VOS models’ performance in real scenarios. Based on LVOS, we evaluate 15 existing VOS models under 3 different settings and conduct a comprehensive analysis. On LVOS, these models suffer a large performance drop, highlighting the challenge of achieving precise tracking and segmentation in real-world scenarios. Attribute-based analysis indicates that one of the significant factors contributing to accuracy decline is the increased video length, interacting with complex challenges such as long-term reappearance, cross-temporal confusion, and occlusion, which emphasize LVOS’s crucial role. We hope our LVOS can advance development of VOS in real scenes.',
  'Lingyi Hong, Zhongying Liu, Wenchao Chen, Chenzhi Tan, Yuang Feng, Xinyu Zhou, Pinxue Guo, Jinglun Li, Zhaoyu Chen, Shuyong Gao, Wei Zhang, Wenqiang Zhang',
  '3',
  '2026',
  '12'],
 ['SUIT: Spatial-Spectral Union-Intersection Interaction Network for Hyperspectral Object Tracking',
  'http://ieeexplore.ieee.org/document/11267013',
  'Hyperspectral videos (HSVs), with their inherent spatial-spectral-temporal structure, offer distinct advantages in challenging tracking scenarios such as cluttered backgrounds and small objects. However, existing methods primarily focus on spatial interactions between the template and search regions, often overlooking spectral interactions, leading to suboptimal performance. To address this issue, this paper investigates spectral interactions from both the architectural and training perspectives. At the architectural level, we first establish band-wise long-range spatial relationships between the template and search regions using Transformers. We then model spectral interactions using the inclusion-exclusion principle from set theory, treating them as the union of spatial interactions across all bands. This enables the effective integration of both shared and band-specific spatial cues. At the training level, we introduce a spectral loss to enforce material distribution alignment between the template and predicted regions, enhancing robustness to shape deformation and appearance variations. Extensive experiments demonstrate that our tracker achieves state-

of-the-art tracking performance. The source code, trained models and results will be publicly available via https://github.com/bearshng/suit to support reproducibility',
  'Fengchao Xiong, Zhenxing Wu, Jun Zhou, Sen Jia, Yuntao Qian',
  '4',
  '2025',
  '12'],
 ['Quality-Aware Spatio-Temporal Transformer Network for RGBT Tracking',
  'http://ieeexplore.ieee.org/document/11270003',
  'Transformer-based RGBT tracking has attracted much attention due to the strong modeling capacity of self attention and cross attention mechanisms. These attention mechanisms utilize the correlations among tokens to construct powerful feature representations, but are easily affected by low-quality tokens. To address this issue, we propose a novel Quality-aware Spatio-temporal Transformer Network (QSTNet), which calculates the quality weights of tokens in search regions based on the correlation with multimodal template tokens to suppress the negative effects of low-quality tokens in spatio-temporal feature representations, for robust RGBT tracking. In particular, we argue that the correlation between search tokens of one modality and multimodal template tokens could reflect the quality of these search tokens, and thus design the Quality-aware Token Weighting Module (QTWM) based on the correlation matrix of search and template tokens to suppress the negative effects of low-quality tokens. Specifically, we calculate the difference matrix derived from the attention matrices of the search tokens from both modalities and the multimodal template tokens, and then assign the quality weight for each search token based on the difference matrix, which reflects the relative correlation of search tokens from different modalities to multimodal template tokens. In addition, we propose the Prompt-based Spatio-temporal Encoder Module (PSEM) to utilize spatio-temporal multimodal information while alleviating the impact of low-quality spatio-temporal features. Extensive experiments on four RGBT benchmark datasets demonstrate that the proposed QSTNet exhibits superior performance compared to other state-of-the-art tracking methods. Our code and supplementary video are now available: https://zhaodongah.github.io/QSTNet',
  'Zhaodong Ding, Chenglong Li, Tao Wang, Futian Wang',
  '4',
  '2025',
  '12'],
 ['Self-Adaptive Vision-Language Tracking With Context Prompting',
  'http://ieeexplore.ieee.org/document/11284903',
  'Due to the substantial gap between vision and language modalities, along with the mismatch problem between fixed language descriptions and dynamic visual information, existing vision-language tracking methods exhibit performance on par with or slightly worse than vision-only tracking. Effectively exploiting the rich semantics of language to enhance tracking robustness remains an open challenge. To address these issues, we propose a self-adaptive vision-language tracking framework that leverages the pre-trained multi-modal CLIP model to obtain well-aligned visual-language representations. A novel context-aware prompting mechanism is introduced to dynamically adapt linguistic cues based on the evolving visual context during tracking. Specifically, our context prompter extracts dynamic visual features from the current search image and integrates them into the text encoding process, enabling self-updating language embeddings. Furthermore, our framework employs a unified one-stream Transformer architecture, supporting joint training for both vision-only and vision-language tracking scenarios. Our method not only bridges the modality gap but also enhances robustness by allowing language features to evolve with visual context. Extensive experiments on four vision-language tracking benchmarks demonstrate that our method effectively leverages the advantages of language to enhance visual tracking. Our large model can obtain 55.0% AUC on $\\text {LaSOT}_{\\text {EXT}}$ and 69.0% AUC on TNL2K. Additionally, our language-only tracking model achieves performance comparable to that of state-of-the-art vision-only tracking methods on TNL2K. Code is available at https://github.com/zj5559/SAVLT',
  'Jie Zhao, Xin Chen, Shengming Li, Chunjuan Bo, Dong Wang, Huchuan Lu',
  '4',

```
  '2025',
  '12'],
 ['CRL: An Efficient Autonomous Exploration Framework for Large-Scale
Environments With Contrastive-Driven Reinforcement Learning',
  'http://ieeexplore.ieee.org/document/11142471',
  'Autonomous exploration in large-scale environments is impeded by two critical
challenges, namely, suboptimal viewpoint selection resulting from inadequate
feature extraction and the continuously rising computational costs as the
environment expands. Existing methods struggle to simultaneously tackle these
dual challenges within cohesive frameworks. In response, we present an efficient
autonomous exploration framework with contrastive-driven reinforcement learning.
Inspired by human cognitive mechanisms that reinforce crucial information
recognition through contrast, our study implements contrastive constraints on
nodes of varying utility levels within high-dimensional feature spaces,
achieving a decoupling of their latent representations. This capability empowers
decision networks to explicitly capture key regional characteristics, thereby
enhancing the precision of optimal viewpoint selection. Moreover, to mitigate
the issues of backtracking and redundant exploration, we design specialized
training rules that enforce effective action constraints, further enhancing
viewpoint selection. Additionally, we propose a novel graph rarefaction
algorithm to tackle computational costs, simplifying computational complexities
while maintaining performance standards. Compared to the state-of-the-art (SOTA)
approaches, our method achieves 6.7% shorter path lengths, while also
demonstrates robust generalization capabilities through real-world robotic
experiments across multiple real-world scenarios.',
  'Benke Gao, Hao Chen, Quan Liu, Hanqiang Deng, Jian Huang, Yan-Jun Liu',
  '5',
  '2025',
  '12'],
 ['End-To-End Multiple Object Detection and Tracking With Spatio-Temporal
Transformers',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/cvi2.70052?af=R',
  'We introduce a novel end-to-end network called the spatio-temporal multiple
object tracking with transformer (STMOTR), which employs motion constraints to
establish binary matching within the reconstructed deformable-DETR network,
heuristically learning object trajectories from the Video Swin backbone. This
subtly constrained matching rule not only keeps the detection ID consistency but
also significantly reduces the potential for tracking ID
switch.\n\n\n\n\n\nABSTRACT\nOptimising both trajectory position information
and identity information is a key challenge in multiple object tracking.
Mainstream approaches ensure ID consistency by combining detection data with
various additional information. However, many methods overlook the inherent
spatio-temporal correlation of trajectory position information. We argue that
additional modules are redundant, and that forecasting trajectories directly
without the need for interframe association by utilising motion constraints is
adequate. In this study, we introduce a novel end-to-end network called the
spatio-temporal multiple object tracking with transformer (STMOTR), which
employs motion constraints to establish binary matching within the reconstructed
deformable-DETR network, heuristically learning object trajectories from the
Video Swin backbone. This subtly constrained matching rule not only keeps the
detection ID consistency but also significantly reduces the potential for
tracking ID switch. We evaluated STMOTR on the UA-DETRAC and our proposed tunnel
multiple object tracking dataset (T-MOT), achieving state-of-the-art performance
with 39.8% PR-MOTA on the UA-DETRAC and 79.6% MOTA on the T-MOT. The source code
is also available at https://github.com/Jade-Ray/STMOTR.',
  'Qi Lei, Xiangyu Song, Shijie Sun, Huansheng Song, Lichen Liu, Zhaoyang
Zhang',
  '6',
  '2025',
  '12'],
 ['ShipsMOT: A Comprehensive Benchmark and Framework for Multiobject Tracking of
Ships',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/cvi2.70042?af=R',
  'A new dataset designed for ship detection and tracking and a new tracking
```

framework designed for ship features.\n\n\n\n\n\nABSTRACT\nMultiobject tracking of ships is crucial for various applications, such as maritime security and the development of ship autopilot systems. However, existing ship visual datasets primarily focus on ship detection tasks, lacking a fully open-source dataset for multiobject tracking research. Furthermore, current methods often struggle with extracting appearance features under complex sea conditions, varying scales and different ship types, affecting tracking precision. To address these issues, we propose ShipsMOT, a new benchmark dataset containing 121 video sequences with an average of 15.45\xa0s per sequence, covering 15 distinct ship types and a total of 237,999 annotated bounding boxes. Additionally, we propose JDR-CSTrack, a ship multiobject tracking framework that improves feature extraction at different scales by optimising a joint detection and Re-ID network. JDR-CSTrack utilises the fusion of appearance and motion features for multilevel data association, thereby minimising track loss and ID switches. Experimental results confirm that ShipsMOT can serve as a benchmark for future research in ship multiobject tracking and validate the superiority of the proposed JDR-CSTrack framework. The dataset and code can be found on https://github.com/jpj0916/ShipsMOT.',
  'Fang Luo, Pengju Jiang, George To Sum Ho, Wenjing Zeng',
  '6',
  '2025',
  '10'],
 ['A Review of Multi-Object Tracking in Recent Times',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/cvi2.70010?af=R',
  'This paper discusses many recent deep-learning MOT methods. Moreover, to highlight their contributions, these methods are categorised into four main groups: detection-based, SOT-based, and segmentation-based methods according to the integrated core technologies.\n\n\n\n\n\nABSTRACT\nMulti-object tracking (MOT) is a fundamental problem in computer vision that involves tracing the trajectories of foreground targets throughout a video sequence while establishing correspondences for identical objects across frames. With the advancement of deep learning techniques, methods based on deep learning have significantly improved accuracy and efficiency in MOT. This paper reviews several recent deep learning-based MOT methods and categorises them into three main groups: detection-based, single-object tracking (SOT)-based, and segmentation-based methods, according to their core technologies. Additionally, this paper discusses the metrics and datasets used for evaluating MOT performance, the challenges faced in the field, and future directions for research.',
  'Suya Li, Hengyi Ren, Xin Xie, Ying Cao',
  '6',
  '2025',
  '3'],
 ['A New Large-Scale Dataset for Marine Vessel Re-Identification Based on Swin Transformer Network in Ocean Surveillance Scenario',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/cvi2.70007?af=R',
  'A new large-scale marine vessel dataset with well-annotated vessel orientation, vessel colour, and vessel type labels has been collected and created in a real marine environment for vessel Re-ID research. (2)A side information embedding module is introduced through the learnable embedding layer to encode more kinds of information, including marine vessel orientation, type and colour. (3)A deep neural network framework based on Swin Transformer for marine vessel Re-ID task is proposed to learn and extract discriminative features, and achieves SOTA performance on vessel, vehicle and person Re-ID benchmark datasets.\n\n\n\n\n\nABSTRACT\nIn recent years, there has been an upward trend that marine vessels, an important object category in marine monitoring, have gradually become a research focal point in the field of computer vision, such as detection, tracking, and classification. Among them, marine vessel re-identification (Re-ID) emerges as a significant frontier research topics, which not only faces the dual challenge of huge intra-class and small inter-class differences, but also has complex environmental interference in the port monitoring scenarios. To propel advancements in marine vessel Re-ID technology, SwinTransReID, a framework grounded in the Swin Transformer for marine vessel Re-ID, is introduced. Specifically, the project initially encodes

the triplet images separately as a sequence of blocks and construct a baseline model leveraging the Swin Transformer, achieving better performance on the Re-ID benchmark dataset in comparison to convolution neural network (CNN)-based approaches. And it introduces side information embedding (SIE) to further enhance the robust feature-learning capabilities of Swin Transformer, thus, integrating non-visual cues (orientation and type of vessel) and other auxiliary information (hull colour) through the insertion of learnable embedding modules. Additionally, the project presents VesselReID-1656, the first annotated large-scale benchmark dataset for vessel Re-ID in real-world ocean surveillance, comprising 135,866 images of 1656 vessels along with 5 orientations, 12 types, and 17 colours. The proposed method achieves 87.1% $\\%$ mAP and 96.1% $\\%$ Rank-1 accuracy on the newly-labelled challenging dataset, which surpasses the state-of-the-art (SOTA) method by 1.9% $\\%$ mAP regarding to performance. Moreover, extensive empirical results demonstrate the superiority of the proposed SwinTransReID on the person Market-1501 dataset, vehicle VeRi-776 dataset, and Boat Re-ID vessel dataset.',
  'Zhi Lu, Liguo Sun, Pin Lv, Jiuwu Hao, Bo Tang, Xuanzhen Chen',
  '6',
  '2025',
  '3'],
 ['Unlocking the power of multi-modal fusion in 3D object tracking',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/cvi2.12335?af=R',
  '3D Single Object Tracking plays a vital role in autonomous driving and robotics, yet traditional approaches have predominantly focused on using pure LiDAR-based point cloud data, often neglecting the benefits of integrating image modalities. To address this gap, we propose a novel Multi-modal Image-LiDAR Tracker (MILT) designed to overcome the limitations of single-modality methods by effectively combining RGB and point cloud data. Our key contribution is a dual-branch architecture that separately extracts geometric features from LiDAR and texture features from images. These features are then fused in a BEV perspective to achieve a comprehensive representation of the tracked object. A significant innovation in our approach is the Image-to-LiDAR Adapter module, which transfers the rich feature representation capabilities of the image modality to the 3D tracking task, and the BEV-Fusion module, which facilitates the interactive fusion of geometry and texture features. By validating MILT on public datasets, we demonstrate substantial performance improvements over traditional methods, effectively showcasing the advantages of our multi-modal fusion strategy. This work advances the state-of-the-art in SOT by integrating complementary information from RGB and LiDAR modalities, resulting in enhanced tracking accuracy and robustness.',
  'Yue Hu',
  '6',
  '2025',
  '2'],
 ['Representation alignment contrastive regularisation for multi-object tracking',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/cvi2.12331?af=R',
  "Achieving high-performance in multi-object tracking algorithms heavily relies on modelling spatial-temporal relationships during the data association stage. Mainstream approaches encompass rule-based and deep learning-based methods for spatial-temporal relationship modelling. While the former relies on physical motion laws, offering wider applicability but yielding suboptimal results for complex object movements, the latter, though achieving high-performance, lacks interpretability and involves complex module designs. This work aims to simplify deep learning-based spatial-temporal relationship models and introduce interpretability into features for data association. Specifically, a lightweight single-layer transformer encoder is utilised to model spatial-temporal relationships. To make features more interpretative, two contrastive regularisation losses based on representation alignment are proposed, derived from spatial-temporal consistency rules. By applying weighted summation to affinity matrices, the aligned features can seamlessly integrate into the data association stage of the original tracking workflow. Experimental results showcase that our model enhances the majority of existing tracking networks' performance without excessive complexity, with minimal increase in training

overhead and nearly negligible computational and storage costs.",
  'Shujie Chen, Zhonglin Liu, Jianfeng Dong, Xun Wang, Di Zhou',
  '6',
  '2025',
  '2'],
 ['Droplet Detection and Tracking in Complex Motions Based on YOLOv5s Network',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ipr2.70245?af=R',
  'This study presents an enhanced model for droplet detection and tracking in solvent extraction operations, using a refined YOLOv5s framework integrated with the OC-SORT algorithm. The model accurately detects droplets and tracks their nonlinear motion trajectories, achieving a detection accuracy of 0.977 and a 40% increase in speed compared to the original model.\n\n\n\n\n\n\n\nABSTRACT\nIn the realm of chemical engineering, solvent extraction is pivotal for separating valuable components from mixtures. Precise understanding of droplet dynamics is crucial for enhancing extraction column efficiency. However, analysing dispersed-phase droplets during solvent extraction operations remains challenging due to their complex motions and interactions. This study proposes an enhanced model for droplet detection and tracking in complex motions. Building upon the YOLOv5s object detection framework, we refine the model to accurately detect droplets and recognise their size characteristics. Additionally, we integrate YOLOv5s with the OC-SORT algorithm to develop a robust droplet tracking model capable of tracing nonlinear droplet motion trajectories under both pulsed and non-pulsed conditions. Experimental results demonstrate that our approach achieves high accuracy and faster detection speeds, with a detection accuracy of 0.977 and a 40% increase in speed compared to the original model. This study not only validates the effectiveness of our model through manual labelling but also provides a solid foundation for further research into the complex fluid dynamics in extraction columns. The source code of the proposed method is available at https://github.com/polangjushi/DDTA.',
  'Yubo Zhang, Zhenning Su, Yong Wang, Boren Tan, Chen Zhao',
  '7',
  '2025',
  '11'],
 ['AI-Powered Human Activity Detection and Tracking in Dense Crowds Using YOLOv8-DeepSORT',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ipr2.70227?af=R',
  'This study presents a real-time human activity detection and recognition system that combines YOLOv8l with deep simple online real-time tracking for enhanced performance in crowded and dynamic environments. Evaluated on over 10,000 frames from public and experimental datasets, the system achieves a mean average precision of 96.1%, with the Adam optimiser outperforming stochastic gradient descent. The model demonstrates robustness and scalability, offering a significant advancement in intelligent public surveillance.\n\n\n\n\n\n\n\nABSTRACT\nMany organisations face challenges in accurately and efficiently identifying human activities using manually operated CCTV cameras and small datasets, particularly in densely populated and dynamic environments. This paper presents a real-time human activity detection and recognition system that leverages You Only Look Once version 8 (YOLOv8) for object detection and integrates the deep simple online real-time tracking algorithm for associate frames, robust tracking and individual identity preservation. Evaluated on a combination of public and experimental datasets containing over 10,000 frames, the system achieves notable enhancements in efficiency along with accuracy. Through data preprocessing and algorithm optimisation, the Adam optimiser outperformed stochastic gradient descent (SGD) by more than 5% in accuracy. K-fold cross-validation was applied to reduce overfitting, while anchor boxes of varying scales and aspect ratios improved the detection of objects of different sizes. The aspect ratio, which is used to mitigate the effect of object size and control edge-orientation issues by computing the ratio of width to height of the bounding boxes in YOLOv8, plays a crucial role in object detection.YOLOv8l, combined with DeepSORT, delivers state-of-the-art results, achieving a mean average precision of 96.1% in mixed datasets. The system assigns unique IDs to track individuals, enabling precise frame association during motion in crowded areas, which is crucial for maintaining accuracy in complex environments. This study presents a robust and

scalable model, offering a significant improvement for public surveillance technology.',
  'Abebe Belay Adege, Muluye Fentie Admas, Yitayal Tehone, Boaz Berhanu Tulu, Eyosiyas Tibebu Endalamaw',
  '7',
  '2025',
  '10'],
 ['Real-Time Underwater Vision Sensing System for AUV Tracking',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ipr2.70183?af=R',
  'In this article, we developed an underwater real-time visual sensing system for AUV tracking based on the edge embedded-device NVIDIA Jetson Xavier NX. After experimental verification, compared with other algorithms, the UW-Siam tracking algorithm we proposed not only has a higher success rate and accuracy, but also has the best speed performance, and can easily cope with underwater object tracking tasks in different scenarios.\n\n\n\n\n\n\n\nABSTRACT\nWith the continuous development of emerging technologies such as big data and artificial intelligence, the related technologies of perception tasks on which underwater object tracking relies have made great progress. However, a significant barrier still exists in implementing high-performance real-time underwater object tracking on low-power edge devices. To achieve real-time tracking of underwater objects for edge devices, this article develops an underwater real-time visual sensing system applied to AUV tracking. First, an underwater object tracking device is designed in this article employing a stereo binocular camera, an edge embedded NVIDIA Jetson Xavier NX and an STM32 control board. Then, after preprocessing, the input image with the effective USM algorithm, we propose a quick approach for detecting underwater objects based on SIoU-YOLOv8n, which enables automatic object recognition and selection. At the same time, this article proposes a twin network UW-Siam for continuous tracking of underwater objects, which achieves more accurate underwater object tracking. Finally, the algorithm is deployed to the designed real-time underwater vision sensing system and tested in real-world scenarios. The tracking accuracy reached 0.652, and the detection mAP reached 0.97. The results indicate that the system can rapidly detect and continuously monitor objects, performing well in real-world scenarios with high accuracy and\xa0robustness.',
  'Canrong Chen, Tingzhuang Liu, Linglu He, Yi Zhu, Fei Yuan',
  '7',
  '2025',
  '8'],
 ['Spatiotemporal Context Adapting Framework for Visual Object Tracking',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ipr2.70150?af=R',
  'Due to noise interference, existing methods often partially exploit auxiliary data, leading to underutilisation of spatiotemporal information. To address these issues, we propose a novel and concise tracking framework, uniformly encoding all auxiliary data and mitigating the unstable factors.\n\n\n\n\n\n\n\nABSTRACT\nVisual object tracking is widely applied in intelligent transportation systems and visual surveillance systems that serve smart cities, as well as in autonomous vehicles. Existing methods usually utilise a relation-modelling framework to model the visual object tracking problem, with auxiliary spatial context and temporal information. The spatial context is often extracted by enlarging the target template, which can introduce more background and positional information. The temporal correlation is obtained by associating the search image with previous images. However, due to noise interference, existing methods often partially exploit auxiliary data, leading to underutilisation of spatiotemporal information. To address these issues, we propose a novel and concise tracking framework, uniformly encoding all auxiliary data, including the enlarged target template, previous images, and corresponding target bounding boxes. Specifically, to mitigate the unstable factors introduced by these raw inputs, we propose a spatiotemporal context adaptive encoder, which can adaptively select appropriate information in noisy data. Extensive experiments show that the proposed method achieves state-of-the-art performance on various benchmarks, demonstrating its\xa0superiority.',
  'Kunlong Zhao, Dawei Zhao, Xu Wang, Liang Xiao, Yulong Huang, Yiming Nie, Yonggang Zhang, Bin Dai',
  '7',

'2025',
'7'],
['A Two-Stage Homography Matrix Prediction Approach for Trajectory Generation in Multi-Object Tracking on Sports Fields',
'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ipr2.70136?af=R',
'To overcome the limitations of existing homography matrix prediction methods, we developed a two-stage prediction approach by optimizing the pixel position feature extraction structure. In the first stage, 4-key-point coordinate deviation is leveraged to indirectly estimate the macroscopic values of the homography matrix, followed by direct fine-tuning of the microscopic 8-DOF numerical components in the second stage. We validated the effectiveness of this approach in three competitive sports scenarios: basketball, ice hockey, and handball.\n\n\n\n\n\n\n\nABSTRACT\nHomography estimation is a fundamental topic in computer vision, especially in scenarios that require perspective changes for intelligent analysis of sports fields, where it plays a crucial role. Existing methods predict the homography matrix either indirectly by evaluating the 4-key-point coordinate deviation in paired images with the same visual content or directly by fine-tuning the 8 degrees of freedom numerical values that define the matrix. However, these approaches often fail to effectively incorporate coordinate positional information and overlook optimal application scenarios, leading to significant accuracy bottlenecks, particularly for paired images with differing visual content. To address these issues, we propose an approach that integrates both methods in a staged manner, leveraging their respective advantages. In the first stage, positional information is embedded to enhance convolutional computations, replacing serial concatenation in traditional feature fusion with parallel concatenation, while using 4-key-point coordinate deviation to predict the macroscopic homography matrix. In the second stage, positional information is further integrated into the input images to refine the direct 8 degrees of freedom numerical predictions, improving matrix fine-tuning accuracy. Comparative experiments with state-of-the-art methods demonstrate that our approach achieves superior performance, yielding a root mean square error as low as 1.25 and an average corner errror as low as 14.1 in homography transformation of competitive sports image\xa0pairs.',
'Pan Zhang, Jiangtao Luo, Guoliang Xu, Xupeng Liang',
'7',
'2025',
'6'],
['Vorticity Transport Equation-Based Shadow Removal Approach for Image Inpainting',
'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ipr2.70114?af=R',
'(1) A new image inpainting algorithm using the VTE is developed; (2) FMM is introduced as an alternative to the iterative numerical solution of the VTE to improve the speed of image in painting; and (3) the VTE equation is transformed into a weighted-average equation to be used in conjunction with FMM.\n\n\n\n\n\n\n\nABSTRACT\nShadows are common in many types of images, causing information loss or disturbance. Shadow removal can help improve the quality of the digital image. If there is no effective information available to restore the original image in the shaded area, the interpolation-based inpainting technique can be used to remove the shadow from the digital image. This image inpainting technique typically involves establishing and solving partial differential equations (PDEs), an iterative solving process that is very time-consuming. To solve the time-consuming problem, a method that introduces the fast marching method (FMM) into the vorticity transport equation (VTE) is demonstrated. VTE is a type of partial differential equation describing two-dimensional fluids. FMM is a numerical scheme for tracking the evolution of monotonically advancing interfaces via finite difference solution of the eikonal equation. The proposed method contains three main steps: (a) by investigating the relationship between VTE and the traditional PDE-based image inpainting method, a new image inpainting model using VTE is developed;(b) the area to be inpainted is divided into boundaries that shrink in layers from the outside inwards using FMM; and (c) the VTE image inpainting model is converted into a weighted average form to coordinate with FMM. The visual and quantitative evaluation of the experimental results of shadow removal shows that the proposed method outperforms PDE-based and state-of-the-art methods in terms of shadow-

removal effect and running time. The results also show that our method excels at inpainting images with near-smooth textures and simple geometric structures and where the pixels to be inpainted are continuous with neighbouring pixels.',
  'Xiaoying Ti, Li Yu, Quanhua Zhao',
  '7',
  '2025',
  '5'],
 ['Structure-Aware Transformer for Shadow Detection',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ipr2.70031?af=R',
  'We propose a structure-aware transformer network for robust shadow detection. An edge-guided multi-task learning framework is designed to predict shadow maps with rich structures. An auxiliary semantic-aware learning is introduced to overcome the interference from complex\xa0scenes.\n\n\n\n\n\n\n\n\nABSTRACT\nShadow detection helps reduce ambiguity in object detection and tracking. However, existing shadow detection methods tend to misidentify complex shadows and their similar patterns, such as soft shadow regions and shadow-like regions, since they treat all cases equally, leading to an incomplete structure of the detected shadow regions. To alleviate this issue, we propose a structure-aware transformer network (STNet) for robust shadow detection. Specifically, we first develop a transformer-based shadow detection network to learn significant contextual information interactions. To this end, a context-aware enhancement (CaE) block is also introduced into the backbone to expand the receptive field, thus enhancing semantic interaction. Then, we design an edge-guided multi-task learning framework to produce intermediate and main predictions with a rich structure. By fusing these two complementary predictions, we can obtain an edge-preserving refined shadow map. Finally, we introduce an auxiliary semantic-aware learning to overcome the interference from complex scenes, which facilitates the model to perceive shadow and non-shadow regions using a semantic affinity loss. By doing these, we can predict high-quality shadow maps in different scenarios. Experimental results demonstrate that our method reduces the balance error rate (BER) by 4.53%, 2.54%, and 3.49% compared to state-of-the-art (SOTA) methods on the benchmark datasets SBU, ISTD, and UCF,\xa0respectively.',
  'Wanlu Sun, Liyun Xiang, Wei Zhao',
  '7',
  '2025',
  '3'],
 ['An intelligent retrievable object-tracking system with real-time edge inference capability',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ipr2.13297?af=R',
  "An intelligent retrievable object-tracking system assists users in quickly and accurately locating lost objects. However, challenges such as real-time processing on edge devices, low image resolution, and small-object detection significantly impact the accuracy and efficiency of video-stream-based systems, especially in indoor home environments. To overcome these limitations, a novel real-time intelligent retrievable object-tracking system is designed. The system incorporates a retrievable object-tracking algorithm that combines DeepSORT and sliding window techniques to enhance tracking capabilities. Additionally, the YOLOv7-small-scale model is proposed for small-object detection, integrating a specialized detection layer and the convolutional batch normalization LeakyReLU spatial-depth convolution module to enhance feature capture for small objects. TensorRT and INT8 quantization are used for inference acceleration on edge devices, doubling the frames per second. Experiments on a Jetson Nano (4 GB) using YOLOv7-small-scale show an 8.9% improvement in recognition accuracy over YOLOv7-tiny in video stream processing. This advancement significantly boosts the system's performance in efficiently and accurately locating lost objects in indoor home\xa0settings.",
  'Yujie Li, Yifu Wang, Zihang Ma, Xinghe Wang, Benying Tan, Shuxue Ding',
  '7',
  '2024',
  '12'],
 ['VT-Surf: Visual Tracking with Switching Dynamics under Time-Series Forecasting',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ell2.70495?af=R',

"We reconceptualize visual tracking as a multivariate time-series forecasting (MTSF) problem. In addition, we propose a principled approach that models the dynamic nature of target motion using a regime-switching framework. This method employs an underlying Markov jump process (MJP) to govern transitions between multiple latent motion patterns, each characterized by its own stochastic differential equation\xa0(SDE).\n\n\n\n\n\n\n\nABSTRACT\nIn this paper, we reconceptualize visual tracking as a multivariate time-series forecasting (MTSF) problem. Specifically, the goal of visual tracking—predicting the target's state over time, including its (x,y)(x, y) center coordinates and scale—can be naturally framed as forecasting future states from a sequence of past observations. Viewed through this lens, visual tracking aligns with the challenges of MTSF, where the objective is to capture complex temporal dependencies among multiple variables. However, applying MTSF to visual tracking introduces new difficulties due to the inherently intricate nature of object motion, which often involves abrupt and nonlinear variations in direction, velocity, and behavioral patterns. To address these complexities, we propose a principled approach that models the dynamic nature of target motion using a regime-switching framework. This method employs an underlying Markov jump process (MJP) to govern transitions between multiple latent motion patterns, each characterized by its own stochastic differential equation\xa0(SDE). By doing so, our model adapts to diverse temporal dynamics in a data-driven manner, enabling robust and precise prediction of future target states. Experimental results demonstrate that our method outperforms state-of-the-art visual tracking approaches, particularly in scenarios where target objects exhibit diverse and dynamic motion patterns over time.",
  'Seonghak Lee, Jisoo Park, Radu Timofte, Junseok Kwon',
  '8',
  '2025',
  '11'],
 ['QMBOC-HFM: Enhanced QMBOC for Large-Scale LEO Navigation Augmentation Systems',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ell2.70398?af=R',
  'A novel QMBOC signal with improved code tracking performance, multipath suppression, and anti-jamming capabilities.\n\n\n\n\n\n\n\nABSTRACT\nLarge-scale low earth orbit (LEO) constellations can enhance traditional global navigation satellite systems (GNSS) in many ways, but the rapid increase in the number of LEO satellites also poses challenges to the anti-interference performance of navigation signals. To improve the performance of the B1C signal in the future BeiDou global navigation satellite system, this paper proposes a modulation method: it enhances the existing quadrature-multiplexed binary offset carrier (QMBOC) by using periodic binary hyperbolic frequency modulation (HFM) signals as subcarriers. The enhanced QMBOC signal (QMBOC-HFM) demonstrates superior anti-interference and positioning accuracy compared to the original QMBOC signal. The proposed QMBOC-HFM provides a potential signal design for future LEO navigation-augmented global satellite navigation systems.',
  'Shugan Zhang, Xinming Huang, Bofang Chen',
  '8',
  '2025',
  '9'],
 ['Maximum Circumnavigation Radius for a Class of Bearings-Only Target Tracking Problem',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ell2.70390?af=R',
  'Under a virtual intersecting localization and tracking algorithm, the maximum circumnavigation radius is proposed for the bearings-only observer by analysing the relationship between the target position estimation error, observation error and the velocity of the observer and the target. The proposed method can provide an engineering application guidance for a class of bearings-only circumnavigation tracking problems. \n\n\n\n\n\n\n\nABSTRACT\nFor the problem of bearings-only target tracking using circumnavigation method, to enhance the safety and stealth, the observer needs to maintain the maximum possible distance from the target. But an excessively large circumnavigation radius may lead to significant tracking error or even no solution in target localization. Under a virtual intersecting localization and tracking algorithm, the maximum circumnavigation radius is proposed for the bearings-only observer by analysing

the relationship between the target position estimation error, observation error and the velocity of the observer and the target, which offers theoretical guidance for the engineering implementation of such circumnavigation tracking problem.',
  'Guoqing Qi, Yinya Li, Andong Sheng',
  '8',
  '2025',
  '8'],
 ['Research on Extended Target Tracking Algorithm Based on RHM and RMM Fusion',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ell2.70371?af=R',
  'This paper addresses the extended target tracking problem with irregular shapes. A fusion algorithm combining random hypersurface model (RHM) and random matrix model (RMM) for extended target tracking is proposed. By leveraging the fast convergence and strong robustness of RMM, the tracking efficiency of RHM is improved. Through simulation experiments, the effectiveness of the algorithm proposed in this paper is verified. The extended target fusion method proposed in this paper can provide a more accurate means for precise tracking and state prediction in practical engineering.\n\n\n\n\n\n\n\n\nABSTRACT\nAiming at the extended target tracking problem with irregular shapes, an extended target tracking algorithm that fuses random hypersurface model (RHM) and random matrix model (RMM) is proposed in the case where the priori shape of the target is unknown. First, the algorithm adopts the RMM method to track the extended target and then takes the reference points on the elliptical shape contour and constructs an ordered sequence. Second, the discrete Fourier transform (DFT) is performed on the sequence to find the coefficients of the harmonic components to obtain the elliptical shape parameters. Then, the ellipsoidal shape parameter is used as the a priori shape parameter of the star-convex RHM. Finally, the effectiveness of the proposed algorithm is verified by simulation experiments. The proposed extended target fusion method provides a more accurate means for precise tracking and state prediction of extended targets.',
  'Guoqing Qi, Zixuan Xu, Yinya Li, Andong Sheng',
  '8',
  '2025',
  '7'],
 ['Longitudinal and Feedforward Controller Design Based on Linear Matrix Inequalities for a Mini Aerial Vehicle',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ell2.70365?af=R',
  'In this paper, the design of a MIMO optimal PID controller and a feedforward controller to enhance disturbance rejection and reference tracking for an MAV system is proposed. The feedback controller employs LMI constraints for iterative variable tuning, while a feedforward strategy is developed to further optimise performance. \n\n\n\n\n\n\n\n\nABSTRACT\nThis paper presents the design of a multi-input multi-output (MIMO) optimal proportional–integral–derivative (PID) controller and a feedforward controller to enhance disturbance rejection and reference tracking in a mini aerial vehicle (MAV). The feedback controller is developed using linear matrix inequality (LMI) constraints within an iterative framework for tuning the PID parameters. In addition, a feedforward control strategy is proposed to further improve performance in reference tracking and disturbance rejection. The feedforward controller is designed under LMI constraints to minimise the H∞ norm of the transfer function matrix relating disturbances (or set-points) to the output (error). This optimisation problem is challenging due to its non-convex and nonlinear nature. Therefore, the main contribution of this paper is to propose a method based on utilising frequency sampling techniques, iterative algorithms and convex optimisation formulation to guarantee optimal solutions for reference tracking and disturbance rejection within specified constraints, which is crucial for preventing irreversible damage in control systems and also enhances the manoeuvrability of the MAV.',
  'Saeedreza Tofighi',
  '8',
  '2025',
  '7'],
 ['FPGA Implementation of a Chopper Controller in PV Systems for IoT Applications',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ell2.70364?af=R',

'A digital control system based on FPGA is developed to generate a chopping frequency ranging from 12.5\xa0kHz to 100\xa0kHz, adjusting its output voltage, frequency, and duty cycle accordingly. The proposed digital controller, with optimisations in architecture and process technology, will improve chip area, decrease delay, and enhance the power efficiency of the converter compared to past researches.\n\n\n\n\n\n\n\n\nABSTRACT\nThis work presents the design and simulation of an Field-Programmable Gate Array (FPGA) based digital controller for a micro-power photovoltaic (PV) energy harvesting system tailored to Internet of Things (IoT) applications. The proposed system features a DC-DC boost converter with an H-bridge topology, designed to interface effectively with low-power PV sources and deliver regulated output to IoT loads. A constant voltage maximum power point tracking (CV-MPPT) algorithm is implemented to maintain optimal energy extraction under variable irradiance. The controller dynamically adjusts the chopping frequency, duty cycle, and output voltage. Leveraging the parallelism and reconfigurability of FPGAs, a high-speed digital control architecture is developed and synthesised on a Cyclone IV FPGA platform. Compared to conventional microcontroller or DSP-based solutions, this approach offers significantly improved response time, power efficiency, and scalability. The final implementation achieves a chip area of 0.92 mm2, a control loop delay of 106\xa0ms, and an overall efficiency of 92.36%, demonstrating its advantages over existing designs in terms of compactness, performance, and energy efficiency.',
'Keyvan Barati, Abdolsamad Hamidi',
'8',
'2025',
'7'],
['High-Impedance Dipoles for Auto-Receiver Tracking Wireless Power Transfer Systems',
'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ell2.70339?af=R',
'This study presents a self-tuning multi-transmitter (multi-Tx) wireless power transfer (WPT) system designed for applications requiring freely positioned receivers (Rxs).\n\n\n\n\n\n\n\n\nABSTRACT\nThis study presents a self-tuning multi-transmitter (multi-Tx) wireless power transfer (WPT) system designed for applications requiring freely positioned receivers (Rxs). The system leverages high-impedance dipole (HID) as Txs to dynamically activate the nearest Tx to a simple dipole antenna as the Rx while deactivating other Txs without additional control circuits. The Txs design incorporates a modified coaxial cable dipole antenna featuring symmetrical gaps over the shield and inductive tuning elements between the central wire and the shield to optimise the resonance frequency and ensure effective power transfer. The feeding point is connected to the central wire via a small cut in the shield. This autonomous behaviour simplifies system operation, enhances flexibility, and ensures acceptable efficient power transfer. Experimental evaluation demonstrates that the system achieves a power transfer efficiency of approximately 65% while maintaining consistent performance as the Rx rotates over two orthogonally located Txs. The dynamic self-tuning mechanism adapts to the position of the Rx, proving particularly advantageous on curved surfaces.',
'Ali Sharifian Mazraeh Mollaei, Ilkka Laakso, Masoud Sharifian Mazraeh Mollaei',
'8',
'2025',
'7'],
['Fine-Grained Temporal Encoding and Decoding-Based Underwater Object Tracking',
'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ell2.70346?af=R',
'This study focuses on this domain and proposes an innovative fine-grained temporal encoding and decoding-based underwater object tracking method.\n\n\n\n\n\n\n\n\nABSTRACT\nUnderwater object tracking is a highly challenging task in the field of computer vision. This study focuses on this domain and proposes an innovative fine-grained temporal encoding and decoding-based underwater object tracking method. Due to the complex and dynamic underwater environment, such as uneven lighting, turbid water quality, and complex target motion patterns, existing underwater object tracking methods face significant limitations in accuracy and stability. By carefully designing a refinement module that combines fine-grained consistency and candidate

elimination, this method can accurately extract fine-grained features of the target and effectively mitigate the interference of various underwater complexities during feature extraction, thereby improving the precision of target features. Furthermore, leveraging the temporal encoding–decoding module, the target features are continuously propagated along the temporal sequence, allowing full utilization of the relational information between frames, which further enhances tracking stability. Experiments were conducted on the UVOT400 dataset, which is large-scale and rich in attributes with diverse target categories. The results demonstrate that, compared to existing methods, this approach significantly outperforms in both accuracy and stability of underwater object tracking, providing new insights and effective solutions for the advancement of underwater object tracking\xa0technology.',
  'Zhen Sun, Zhenggang Guan, Qinghua Li, Mengyang Yuan, Haonan Sun',
  '8',
  '2025',
  '6'],
 ['Full-Space Scanning Leaky-Wave Antenna Utilising Spoof Surface Plasmon Polaritons',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ell2.70328?af=R',
  'In this paper, we present a novel spoof surface plasmon polariton leaky-wave antenna (SSPP-LWA) capable of full-space beam scanning. The proposed antenna achieves full-space beam scanning from -90° to +90° across the frequency band of 9.8–13.3\xa0GHz, with a high scanning rate of 5.94° per 1% frequency variation. Additionally, the design features a simplified structure conducive to practical fabrication. \n\n\n\n\n\n\n\nABSTRACT\nA novel spoof surface plasmon polariton leaky-wave antenna (SSPP-LWA) capable of full-space beam scanning is proposed in this letter. First, the surface plasmon polariton (SPPs) unit and its corresponding transmission line are carefully designed, and the underlying mechanism to achieve full-space scanning is analysed based on a monopole array structure. Specifically, a spoof surface plasmon polariton (SSPP) transmission line combined with a semicircular, three-dimensionally loaded radiating patch is introduced and equivalently modelled as a monopole radiating element. Subsequently, the proposed LWA is realised by employing a monopole array composed of 12 identical radiating elements. Simulation results verify that the antenna achieves continuous beam scanning from -90° to +90° within the frequency range of 9.8–13.3\xa0GHz, exhibiting a high scanning rate of up to 5.94°/%, with an average gain of approximately 9 dBi. Benefiting from its excellent scanning capability and stable radiation characteristics, this antenna demonstrates significant potential for target detection and tracking applications in radar systems.',
  'Ran Yu, Leilei Liu',
  '8',
  '2025',
  '6'],
 ['Implementation and Performance Analysis of RTK-GNSS in Wearable Devices for Athletes in Harsh Environments',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ell2.70289?af=R',
  'This study investigates the implementation of real-time kinematic global navigation satellite system (RTK-GNSS) technique in wearable electronic performance and tracking system for football players to obtain highly accurate positioning data, even in harsh environments. The experiment results demonstrate that the utilization of RTK-GNSS can significantly improve positioning accuracy in various stadium environments, which is crucial for evaluating player performance.\n\n\n\n\n\n\n\nABSTRACT\nThis study presents the performance analysis of RTK-GNSS (real-time kinematic global navigation satellite system) mounted on wearable devices for athletes in harsh environments. GNSS signal can deteriorate by obstacles, i.e. stadium roofs, due to the high GDOP (geometric dilution of precision). RTK-GNSS is used to overcome harsh environment problems in this study. Hardware experiments are performed to analyse the performance of RTK-GNSS in sports wearable devices in harsh environments. The full pitch tracking test results of a wearable device equipped with RTK-GNSS demonstrate the best performance among three different wearable devices in a harsh environment; the maximum positioning error of 1.56 m$\\text{m}$ is approximately 3–30 times smaller than that of wearable devices equipped with stand-alone GNSS.

The distance test results of the wearable device equipped with RTK-GNSS validate the most accurate performance, 98.97%$\\%$, compared to the true value in harsh\xa0environment.',
  'Mingu Kim, Bit Kim, Chulwoo Park, Jinsung Yoon',
  '8',
  '2025',
  '5'],
 ['Robust Beam Tracking for 3D Manoeuvrable UAV in DFRC Systems',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ell2.70211?af=R',
  "This letter presents a novel 3D robust manoeuvrable unmanned aerial vehicle (UAV) beam tracking algorithm within the dual-functional radar-communication framework. Specifically, we propose the interacting multiple model algorithm to accommodate multiple potential UAV motion models and improve the unscented Kalman filter algorithm to enhance its stability.\n\n\n\n\n\n\n\n\nABSTRACT\nDual-functional radar-communication (DFRC) will be a key technology in future sixth-generation (6G) network. Specifically, an integrated radar and communication platform (IRCP) equipped with full-dimensional antenna arrays can execute 3D beamforming, which can effectively minimize interference for communicating with unmanned aerial vehicles (UAVs). However, a significant challenge in fully exploiting 3D beamforming gain is the IRCP's ability to precisely track manoeuvrable UAVs. In this letter, we propose a novel interacting multiple model with enhanced unscented Kalman filter algorithm to realize 3D beam tracking for one manoeuvrable UAV within the framework of DFRC. To be specific, we build multiple state transition models to adopt the manoeuvrability of the UAV. Meanwhile, the stability of the unscented Kalman filter is improved by using the singular value decomposition. Simulation results show that the proposed algorithm has higher tracking accuracy and better transmission performance for one manoeuvrable UAV than the traditional single-motion model based beam tracking\xa0algorithm.",
  'Yuhang Tang, Wei Liu, Jinkun Zhu, Jing Lei, Haoying Mo',
  '8',
  '2025',
  '3'],
 ['Optimal Performance Analysis for Networked System with Quantitative Control Input Over Feedback Channel',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ell2.70183?af=R',
  "This paper focuses on the performance analysis of networked control systems, specifically addressing the optimal tracking performance with quantitative control inputs over feedback channels. A dual-degree-of-freedom controller is designed using the Youla parameterization method, and the system's performance is assessed through numerical simulations. The findings highlight the impact of unstable poles, non-minimum phase zeros, and quantization errors on tracking performance, offering insights for practical control system design.\n\n\n\n\n\n\n\n\nABSTRACT\nNetworked control systems (NCS) have become an emerging and important research area given the rapidly developing network communication technology. Numerous studies in this area concentrate on the stability of control systems, while little has been done on its performance analysis, especially the index of an optimal control system. We fill this gap here by analysing the optimal tracking performance for an NCS that has quantitative control inputs over feedback channels. We design a dual-degree-of-freedom controller along with an optimal control system, by applying the Youla parameterization method based on the time domain and coprime decompositions. The proposed methodology is assessed by numerical simulations. The system's tracking performance is found to be deteriorated by the unstable poles and non-minimum phase zeros of the controlled object, as well as by quantization signal errors. These findings are beneficial for analysing and designing practical control\xa0systems.",
  'Fang Han, Xiaowei Jiang',
  '8',
  '2025',
  '3'],
 ['Applying Adaptive Wavelet Neural Network and Sliding Mode Control for Tracking Control of MEMS Gyroscope',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ell2.70187?af=R',

'In this letter, an algorithm applying an adaptive wavelet neural network (AWNN) and sliding-mode control (SMC) is proposed, investigated and exploited for tracking control of micro-electromechanical system (MEMS) gyroscope. Such an AWNN model can be regarded as a special radius basis function neural network and utilizes Mexican hat function as activation function. Besides, Taylor expansion is used for analyzing activation radius, which is considered as an adaptive variable. The parameters of the MEMS gyroscope model are hard to obtain in engineering applications; thus, AWNN is designed to approximate the uncertain function of MEMS gyroscope and the unknown asymmetrical dead zone in the control scheme. The weights updating laws and the activation radius adaptive laws in AWNN are derived from the Lyapunov stability analysis, which results in the control error converging to the desired value and the weights and activation radius converging to its real value. To achieve the effect of error acceleration, a power function is used to design a sliding mode function. Computer simulation results substantiate the theoretical analysis and further demonstrate the efficacy of such an algorithm combined with AWNN and SMC for MEMS gyroscope control.',
'Guo Luo, Bingling Chen',
'8',
'2025',
'3'],
['A GGIW PMBM Smoother for Multiple Extended Object Tracking',
'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ell2.70188?af=R',
'A gamma Gaussian inverse Wishart Poisson multi-Bernoulli mixture (PMBM) filter that uses multiple time scans for smoothing is proposed. The results of continuous tracking are smoothed and corrected at regular intervals, resulting in more accurate tracking performance. Simulation results show that the proposed algorithm effectively reduces generalized optimal sub-pattern assignment errors compared to the original PMBM filter. \n\n\n\n\n\n\n\nABSTRACT\nPoisson multi-Bernoulli mixture (PMBM) filtering is an effective method for multi-extended target tracking. However, it has a low utilization of information from multiple time scans, leading to potential errors such as jumps and discontinuities in the target trajectories. To address this, propose a gamma Gaussian inverse Wishart PMBM filter that uses multiple time scans for smoothing. The results of continuous tracking are smoothed and corrected at regular intervals, resulting in more accurate tracking performance. Simulation results show that the proposed algorithm effectively reduces generalized optimal sub-pattern assignment errors compared to the original PMBM filter.',
'Wenhui Wang, Ye Xu, Kejie Zhang, YouPeng Sun, Peng Li',
'8',
'2025',
'2'],
['A Novel D-InSAR and Offset Tracking Fusion Algorithm Based on Variance-Masked Mean Filtering',
'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ell2.70173?af=R',
'This paper proposes a novel fusion algorithm based on variance-masked mean filtering to combine the strengths of both methods. The algorithm filters noise in gentle regions of the offset map and fuses refined results with D-InSAR to enhance accuracy. Applied to the 2023 Türkiye earthquake, the method effectively reduces noise, corrects unwrapping errors, and achieves superior accuracy in high-gradient deformation scenarios such as earthquakes and landslides.\n\n\n\n\n\n\n\nABSTRACT\nAccurate ground deformation monitoring is crucial for understanding geophysical activities and mitigating disaster risks. D-InSAR and offset tracking are widely used but face limitations: D-InSAR struggles with phase unwrapping errors in large-scale deformation regions, while offset tracking lacks precision in small-scale deformation. This paper proposes a fusion algorithm based on variance-masked mean filtering to combine the strengths of both methods. The algorithm filters noise in gentle regions of the offset map and fuses refined results with D-InSAR to enhance accuracy. Applied to the 2023 Türkiye earthquake, the method effectively reduces noise, corrects unwrapping errors, and achieves superior accuracy in large-scale deformation scenarios such as earthquakes and landslides.',
'Yifan Wang, Gen Li, Zihan Hu, Chuanxu Sun, Zhe Li',
'8',

```
  '2025',
  '2'],
 ['Extended target tracking using neural network and Gaussian process',
  'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ell2.70151?af=R',
  "In extended target tracking, Gaussian Process (GP) is utilized to model
unknown contour functions based on the model-predicted target center and contour
measurements. However, model prediction relies on accurate prior knowledge. When
the model-predicted target center is inaccurate, it will affect the modelling of
the measurement model. To address issue, this letter introduces a hybrid-driven
approach that combines extended Kalman filter using GP with neural network;
proposes an extended target tracking algorithm using neural network and GP. The
algorithm predicts the target center according to the neural network and the
target's kinematic model, and takes the prediction center and the contour
measurements at the current moment as the input of the neural network, which in
turn provides real-time estimates for the predicted center compensation. The
simulation results show that the algorithm has a significant improvement in
tracking performance and better accuracy in estimating the center position and
extent state of the target.",
  'Hao Wang, Liping Song',
  '8',
  '2025',
  '1']]

[['Curvi-Tracker: Curvilinear structure segmentation refinement by iterative
tracking',
  'https://www.sciencedirect.com/science/article/pii/S0031320325014608',
  '',
  'Zhan Heng, Maurice Pagnucco, Erik Meijering, Yang Song',
  '2',
  '2026',
  '5'],
 ['SNNTracker: Online High-Speed Multi-Object Tracking With Spike Camera',
  'http://ieeexplore.ieee.org/document/11165142',
  'Multi-object tracking (MOT) is crucial for applications such as autonomous
driving and robotics, yet traditional image-based methods struggle in high-speed
scenarios due to motion blur and temporal gaps caused by low frame rates. Spike
cameras, with their ability to continuously record spatiotemporal signals,
overcome these limitations. However, existing spike-based methods often rely on
intermediate image reconstruction or discrete clustering, limiting real-time
performance and temporal continuity. To address this, we propose SNNTracker, the
first fully spiking neural network (SNN)-based MOT algorithm tailored for spike
cameras. SNNTracker integrates a dynamic neural field (DNF)-based attention
mechanism for target detection and a winner-take-all (WTA)-based tracking module
with online spike-timing-dependent plasticity (STDP) for adaptive learning of
object trajectories. By directly processing spike streams without
reconstruction, SNNTracker reduces latency, computational overhead, and
dependency on image quality, making it ideal for ultra-high-speed environments.
It maintains robust, continuous tracking even under occlusions, severe lighting
variations, or temporary object disappearance, by leveraging SNN-estimated
motion predictions and long-term online clustering. We construct three types of
spike-camera MOT datasets covering dense and sparse annotations across diverse
real-world scenarios, including camera ego-motion, deformable and ultra-fast
motion (up to 2600 RPM), occlusion, indoor/outdoor lighting changes, and low-
visibility tracking. Extensive experiments demonstrate that SNNTracker
consistently outperforms state-of-the-art MOT methods—both ANN- and SNN-based—
achieving MOTA scores above 96% and up to 100% in many sequences. Our results
highlight the advantages of spike-driven SNNs for low-latency, high-speed, and
label-free multi-object tracking, advancing neuromorphic vision for real-time
perception.',
  'Yajing Zheng, Chengen Li, Jiyuan Zhang, Zhaofei Yu, Tiejun Huang',
  '3',
  '2026',
  '12'],
```

['SUIT: Spatial-Spectral Union-Intersection Interaction Network for Hyperspectral Object Tracking',
'http://ieeexplore.ieee.org/document/11267013',
'Hyperspectral videos (HSVs), with their inherent spatial-spectral-temporal structure, offer distinct advantages in challenging tracking scenarios such as cluttered backgrounds and small objects. However, existing methods primarily focus on spatial interactions between the template and search regions, often overlooking spectral interactions, leading to suboptimal performance. To address this issue, this paper investigates spectral interactions from both the architectural and training perspectives. At the architectural level, we first establish band-wise long-range spatial relationships between the template and search regions using Transformers. We then model spectral interactions using the inclusion-exclusion principle from set theory, treating them as the union of spatial interactions across all bands. This enables the effective integration of both shared and band-specific spatial cues. At the training level, we introduce a spectral loss to enforce material distribution alignment between the template and predicted regions, enhancing robustness to shape deformation and appearance variations. Extensive experiments demonstrate that our tracker achieves state-of-the-art tracking performance. The source code, trained models and results will be publicly available via https://github.com/bearshng/suit to support reproducibility',
'Fengchao Xiong, Zhenxing Wu, Jun Zhou, Sen Jia, Yuntao Qian',
'4',
'2025',
'12'],
['Unlocking the power of multi-modal fusion in 3D object tracking',
'https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/cvi2.12335?af=R',
'3D Single Object Tracking plays a vital role in autonomous driving and robotics, yet traditional approaches have predominantly focused on using pure LiDAR-based point cloud data, often neglecting the benefits of integrating image modalities. To address this gap, we propose a novel Multi-modal Image-LiDAR Tracker (MILT) designed to overcome the limitations of single-modality methods by effectively combining RGB and point cloud data. Our key contribution is a dual-branch architecture that separately extracts geometric features from LiDAR and texture features from images. These features are then fused in a BEV perspective to achieve a comprehensive representation of the tracked object. A significant innovation in our approach is the Image-to-LiDAR Adapter module, which transfers the rich feature representation capabilities of the image modality to the 3D tracking task, and the BEV-Fusion module, which facilitates the interactive fusion of geometry and texture features. By validating MILT on public datasets, we demonstrate substantial performance improvements over traditional methods, effectively showcasing the advantages of our multi-modal fusion strategy. This work advances the state-of-the-art in SOT by integrating complementary information from RGB and LiDAR modalities, resulting in enhanced tracking accuracy and robustness.',
'Yue Hu',
'6',
'2025',
'2']]