



Intelligence & Robotics

Articles Collection (2024)

Dear Readers of *Intelligence & Robotics*,

Here is a collection of published articles in *Intelligence & Robotics (IR)* in 2024, including title, author, citation format, abstract, and keywords.

You may click titles in CONTENTS or search keywords to locate the article quickly, then you can read and download the articles you picked for free. You can also click Export Citation on the article page to import it into your Endnote. We hope this collection can bring some convenience to your reading and research.

If you have any questions or suggestions on this collection, please contact us. The Editorial Office of *IR* would be most delighted to hear from you.

Editorial Office

Intelligence & Robotics

Website: <https://www.oaepublish.com/ir>

Online Submission: <https://www.oaecenter.com/login?JournalId=ir>

CONTENTS

Editorial	1
1. Embracing the future: the rise of humanoid robots and embodied AI.....	1
Review	1
1. A survey of datasets in medicine for large language models	1
2. Distributed model predictive control for unmanned aerial vehicles and vehicle platoon systems: a review	1
Research Article	2
1. Infrared and visible image fusion based on multi-level detail enhancement and generative adversarial network	2
2. SANet: scale-adaptive network for lightweight salient object detection	3
3. Time-optimal trajectory planning for a six-degree-of-freedom manipulator: a method integrating RRT and chaotic PSO	3
4. Dynamic event-triggered integral non-singular terminal sliding mode tracking of unmanned surface vehicle via an event-triggered extended state observer and adaptive neural network	4
5. An automotive tire visual laser marking robot system based on multi-information fusion.....	5
6. VNN-DM: a vector neural network-based detection model for time synchronization attacks in park-level energy internet	5
7. Path planning method for USVs based on improved DWA and COLREGs	6
8. Improved DDPG algorithm-based path planning for unmanned surface vehicles	7
9. Neurodynamics-based formation tracking control of leader-follower nonholonomic multiagent systems	7
10. Intelligent prediction of rail corrugation evolution trend based on self-attention bidirectional TCN and GRU	8
11. An in-vehicle real-time infrared object detection system based on deep learning with resource-constrained hardware	8
12. Parallel implementation for real-time visual SLAM systems based on heterogeneous computing	9
13. Importance-driven denial-of-service attack strategy design against remote state estimation in multi-agent intelligent power systems	10
14. A novel fatigue driving detection method based on whale optimization and Attention-enhanced GRU	10
15. Parameter identification of an open-frame underwater vehicle based on numerical simulation and quantum particle swarm optimization	11
16. Structural damage identification method based on Swin Transformer and continuous wavelet transform	12
17. Towards environment perception for walking-aid robots: an improved staircase shape feature extraction method.....	12
18. A deep learning-based system for accurate detection of anatomical landmarks in colon environment	13
19. Adaptive prescribed performance tracking control for underactuated unmanned surface ships with input quantization.....	13
20. A novel zero-force control framework for post-stroke rehabilitation training based on fuzzy-PID method	14

21. Coordinated energy-efficient walking assistance for paraplegic patients by using the exoskeleton-walker system	14
22. Leveraging active queries in collaborative robotic mission planning	15
23. A small-sample time-series signal augmentation and analysis method for quantitative assessment of bradykinesia in Parkinson's disease	16
24. Metropolis criterion pigeon-inspired optimization for multi-UAV swarm controller.....	16
25. A collaborative siege method of multiple unmanned vehicles based on reinforcement learning	17
26. Pig-ear detection from the thermal infrared image based on improved YOLOv8n.....	18
27. Recognition of the behaviors of dairy cows by an improved YOLO	18

Editorial

1. Embracing the future: the rise of humanoid robots and embodied AI

Authors: Jiankun Wang*, Chaoqun Wang, Weinan Chen, Qi Dou, Wenzheng Chi

How to cite: Wang, J.; Wang, C.; Chen, W.; Dou, Q.; Chi, W. Embracing the future: the rise of humanoid robots and embodied AI. *Intell. Robot.* **2024**, *4*, 196-9. <http://dx.doi.org/10.20517/ir.2024.12>

Review

1. A survey of datasets in medicine for large language models

Authors: Deshiwei Zhang, Xiaojuan Xue, Peng Gao, Zhijuan Jin, Menghan Hu*, Yue Wu, Xiayang Ying*

How to cite: Zhang, D.; Xue, X.; Gao, P.; Jin, Z.; Hu, M.; Wu, Y.; Ying, X. A survey of datasets in medicine for large language models. *Intell. Robot.* **2024**, *4*, 457-78. <http://dx.doi.org/10.20517/ir.2024.27>

Abstract: With the advent of models such as ChatGPT and other models, large language models (LLMs) have demonstrated unprecedented capabilities in understanding and generating natural language, presenting novel opportunities and challenges within the medicine domain. While there have been many studies focusing on the employment of LLMs in medicine, comprehensive reviews of the datasets utilized in this field remain scarce. This survey seeks to address this gap by providing a comprehensive overview of the datasets in medicine fueling LLMs, highlighting their unique characteristics and the critical roles they play at different stages of LLMs' development: pre-training, fine-tuning, and evaluation. Ultimately, this survey aims to underline the significance of datasets in realizing the full potential of LLMs to innovate and improve healthcare outcomes.

Keywords: Large language models (LLMs), NLP, dataset in medicine, Q&A system in medicine

2. Distributed model predictive control for unmanned aerial vehicles and vehicle platoon systems: a review

Authors: Yang Peng, Huaicheng Yan*, Kai Rao, Penghui Yang, Yunkai Lv

How to cite: Peng, Y.; Yan, H.; Rao, K.; Yang, P.; Lv, Y. Distributed model predictive control for unmanned aerial vehicles and vehicle platoon systems: a review. *Intell. Robot.* **2024**, *4*, 293-317. <http://dx.doi.org/10.20517/ir.2024.19>

Abstract: This paper reviews the application of distributed model predictive control (DMPC) for autonomous intelligent systems (AIS) with unmanned aerial vehicles (UAVs) and vehicle platoon systems. DMPC is an optimal control method that formulates and solves optimization problems to

adjust control strategies by predicting future states based on system models while managing constraints, and this technique has been applied to an increasing number of industrial areas. As the essential parts of AIS, UAVs and vehicle platoon systems have received extensive attention in the civil, industrial, and military fields. DMPC has the ability to quickly solve optimization problems in real-time while taking into account the prediction of the future state of the system, which fits in well with the ability of AIS to predict the environment when making decisions, so the application of DMPC in AIS has a natural advantage. This paper first introduces the basic principles of DMPC and the theoretical results in multi-agent systems (MASs). It then reviews the application of DMPC methods to UAVs and vehicle platoon systems. Finally, the challenges of the existing methods are summarized to offer insights to advance the future development of DMPC in practical applications.

Keywords: Distributed model predictive control, autonomous intelligent systems, multi-agent systems, unmanned aerial vehicles, vehicle platoon systems

Research Article

1. Infrared and visible image fusion based on multi-level detail enhancement and generative adversarial network

Authors: Xiangrui Tian*, Xiaohan Xianyu, Zhimin Li, Tong Xu, Yinjun Jia

How to Cite: Tian, X.; Xianyu, X.; Li, Z.; Xu, T.; Jia, Y. Infrared and visible image fusion based on multi-level detail enhancement and generative adversarial network. *Intell. Robot.* **2024**, *4*, 524-43. <http://dx.doi.org/10.20517/ir.2024.30>

Abstract: Infrared and visible image fusion technology has a wide range of applications in many fields such as target detection and tracking. Existing image fusion methods often overlook the scale hierarchical structure information of features, with local and global features not being closely interconnected. Typically, improvements focus on the network structure and loss functions, while the intimate connection between the quality of the source images and the feature extraction network is often neglected. The aforementioned issues lead to artifacts and blurring of fused images; besides, the detailed edge information can not be well reflected. Therefore, a method of infrared and visible image fusion based on a generative adversarial network (GAN) with multi-level detail enhancement is proposed in this paper. Firstly, the edge information of the input source image is enriched by the multi-level detail enhancement method, which improves the image quality and makes it more conducive to the learning of feature extraction network. Secondly, the residual-dense and multi-scale modules are designed in the generator and the connection between local and global features is established to ensure the transmissibility and coherence of the feature information. Finally, by designing the loss function and dual discriminator constraints to constrain the fusion image, more structure and detail information are added in continuous confrontation. The experimental results show that the fused image contains more detailed texture information and prominent thermal radiation targets. It also outperforms other fusion methods in terms of average gradient (AG), spatial frequency (SF) and edge intensity (EI) metrics, with values surpassing the sub-optimal metrics of 65.41%, 65.09% and 55.22%, respectively.

Keywords: Image fusion, multi-level detail enhancement, generative adversarial networks, deep feature extraction

2. SANet: scale-adaptive network for lightweight salient object detection

Authors: Zhuang Liu, Weidong Zhao*, Ning Jia, Xianhui Liu, Jiexiong Yang

How to Cite: Liu, Z.; Zhao, W.; Jia, N.; Liu, X.; Yang, J. SANet: scale-adaptive network for lightweight salient object detection. *Intell. Robot.* **2024**, *4*, 503-23. <http://dx.doi.org/10.20517/ir.2024.29>

Abstract: Salient object detection (SOD) is widely used in transportation such as road damage detection, assisted driving, etc. However, heavyweight SOD methods are difficult to apply in scenarios with low computing power due to their huge amount of computation and parameters. The detection accuracy of most lightweight SOD methods is difficult to meet application requirements. We propose a novel lightweight scale-adaptive network to achieve a trade-off between lightweight restriction and detection performance. We first propose the scale-adaptive feature extraction (SAFE) module, which mainly consists of two parts: multi-scale feature interaction, which can extract features of different scales and enhance the representation ability of the network; and dynamic selection, which can adaptively assign different weights to features of varying scales according to their contribution through the input image. Then, based on the SAFE module, a lightweight and adaptive backbone network is designed, and scale-adaptive network is implemented in combination with the multi-scale feature aggregation (MFA) module. We evaluate the model quantitatively and qualitatively on six public datasets and compare it with typical heavyweight and lightweight methods. With only 2.29 M parameters, it can achieve a prediction speed of 62 fps on a GTX 3090 GPU, far exceeding other models, and real-time performance is guaranteed. The model performance reaches that of general heavyweight methods and exceeds state-of-the-art lightweight methods.

Keywords: Salient object detection, lightweight SOD, model lightweighting, multi-scale learning

3. Time-optimal trajectory planning for a six-degree-of-freedom manipulator: a method integrating RRT and chaotic PSO

Authors: Zuoxun Wang*, Chuanzhe Pang, Jinxue Sui, Guojian Zhao, Wangyao Wu, Liteng Xu

How to Cite: Wang, Z.; Pang, C.; Sui, J.; Zhao, G.; Wu, W.; Xu, L. Time-optimal trajectory planning for a six-degree-of-freedom manipulator: a method integrating RRT and chaotic PSO. *Intell. Robot.* **2024**, *4*, 479-502. <http://dx.doi.org/10.20517/ir.2024.28>

Abstract: This study proposes an innovative algorithm based on a hybrid optimization strategy, integrating the rapidly-exploring random tree (RRT) and an improved particle swarm optimization (PSO) method to address the time-optimal trajectory planning problem for six-degree-of-freedom robotic arms. The proposed approach emphasizes obstacle avoidance and motion smoothness. RRT is utilized within a dynamic three-dimensional environment to rapidly generate an initial collision-

free path. Subsequently, improved PSO enhances global search performance by introducing a multi-source chaotic mapping-based population initialization strategy, dynamically adjusted inertia weights, and nonlinear learning factors. These enhancements effectively mitigate the limitations of traditional PSO methods, which are prone to premature convergence in complex optimization problems. Furthermore, the proposed 3-5-3 polynomial interpolation method significantly smooths the trajectory, reducing fluctuations in velocity and acceleration, and thereby improving the precision and energy efficiency of trajectory planning. Experimental results demonstrate that the proposed algorithm outperforms existing methods, such as the improved RRT and improved non-dominated sorting genetic algorithm, across multiple metrics. Notable achievements include a reduction of total motion time by approximately 21%, improved stability in robotic arm motion, and enhanced adaptability to dynamic environments. Particularly, the method achieves superior trajectory diversity and uniformity through joint optimization of multiple chaotic sources, overcoming the inherent limitations of single chaotic mappings. This research expands the application scenarios of RRT and PSO and provides a novel solution for intelligent control and real-time planning of high-degree-of-freedom robotic arms.

Keywords: Chaotic map, rapid random tree, particle swarm optimization algorithm, 3-5-3 polynomial

4. Dynamic event-triggered integral non-singular terminal sliding mode tracking of unmanned surface vehicle via an event-triggered extended state observer and adaptive neural network

Authors: Xingmin Wang, Ruixue Liu, Aleksander Sladkowski, Qian Li, Ru Jiang*

How to Cite: Wang, X.; Liu, R.; Sladkowski, A.; Li, Q.; Jiang, R. Dynamic event-triggered integral non-singular terminal sliding mode tracking of unmanned surface vehicle via an event-triggered extended state observer and adaptive neural network. *Intell. Robot.* **2024**, *4*, 439-56. <http://dx.doi.org/10.20517/ir.2024.26>

Abstract: This paper presents a novel trajectory tracking controller for an underactuated unmanned surface vehicle (USV). The controller incorporates an event-triggered extended state observer (ETESO), a minimum learning parameter neural network, an integral non-singular terminal sliding mode (INTSM) control strategy, and a dynamic event-triggered mechanism (DETM). Firstly, an ETESO is developed to estimate unmeasurable velocities and lumped disturbances, differentiating it from most existing extended state observers without the necessity for real-time output measurements. To further alleviate the communication burden and minimize actuator wear, a DETM with an adjustable threshold is introduced. In contrast to traditional event-triggered methods, which employ fixed threshold parameters, this mechanism allows for online adaptive updates of the triggering thresholds, thereby enhancing resource efficiency. Additionally, an INTSM is designed to ensure rapid convergence of the position and velocity errors of the USV. To effectively counteract external disturbances and internal modeling uncertainties, a minimum learning parameter (MLP) neural network algorithm is implemented to approximate and compensate for these uncertainties. Finally, using Lyapunov's theory, it is demonstrated that all signals within the closed-loop tracking

control system remain bounded. Simulation results are given to illustrate the effectiveness of theoretical results.

Keywords: Event-triggered extended state observer, dynamic event-triggered mechanism, USV, integral non-singular terminal sliding mode control

5. An automotive tire visual laser marking robot system based on multi-information fusion

Authors: Chuanxiang Ren, Yunrui Xu, Xiang Liu*, Haibo Liu

How to Cite: Ren, C.; Xu, Y.; Liu, X.; Liu, H. An automotive tire visual laser marking robot system based on multi-information fusion. *Intell. Robot.* **2024**, *4*, 422-38. <http://dx.doi.org/10.20517/ir.2024.25>

Abstract: As smart manufacturing technology continues to advance, laser marking robots have been applied to automotive tire marking, revolutionizing the traditional manual process. In order to improve productivity, these robots face challenges posed by tire positioning, environmental variations, and the light-absorbing properties of the material. In order to solve the problem, a robot vision modeling method based on the fusion of the 3D point cloud information and 2D image information on the surface of automotive tires is presented and is used to construct a visual laser marking robot system for automotive tires. The constructed visual laser marking robot system for automotive tires has been tested and the results show that the laser marking is more effective compared to the traditional manual marking process; the laser marking robot system equipped with a multi-information fusion vision model increases the marking success rate by 8%, increases the speed by nearly nine times, reduces the waste tire rate by 8%, and reduces the economic consumption by nearly 56 times; compared with a single vision information marking robot system, the marking success rate increases by 3% and the tire waste rate reduces by 3%.

Keywords: Information fusion, vision modeling, robot, automotive tire, laser marking

6. VNN-DM: a vector neural network-based detection model for time synchronization attacks in park-level energy internet

Authors: Jiacheng Yang, Fanrong Shi*, Yunlong Li, Zhihang Zhao, Qiushi Cui

How to Cite: Yang, J.; Shi, F.; Li, Y.; Zhao, Z.; Cui, Q. VNN-DM: a vector neural network-based detection model for time synchronization attacks in park-level energy internet. *Intell. Robot.* **2024**, *4*, 406-21. <http://dx.doi.org/10.20517/ir.2024.24>

Abstract: Micro phasor measurement units (μ PMUs) provide high-precision voltage and current phasor data, allowing real-time state estimation and fault detection, which are critical for the stability and reliability of modern power systems. However, their reliance on accurate time synchronization makes them vulnerable to time synchronization attacks (TSAs), which can disrupt grid monitoring <https://intellrobot.com/>

and control by corrupting μ PMU data. Addressing these vulnerabilities is essential to ensure the secure and resilient operation of smart grids and energy internet technologies. To address these challenges, intelligent detection methods are essential. Therefore, this paper proposes a μ PMU measurement data TSA detection model based on vector neural networks (VNNs). This model initially employs a vector neural network to process raw data, effectively extracting and analyzing temporal features. During the same time, a capsule network is employed to classify these temporal features. On this basis, a reconstruction network is used to verify the representational capacity of the model. Simulations based on μ PMU measurement data demonstrate that the model exhibits excellent detection capacity in various performance metrics, underscoring its precision and robustness.

Keywords: Time synchronization attack, vector neural networks, μ PMUs, park-level energy internet, attack detection

7. Path planning method for USVs based on improved DWA and COLREGs

Authors: Shiqi Liu, Xingmin Wang, Yang Wu, Qian Li*, Jiuxiang Yan, Eugene Levin

How to Cite: Liu, S.; Wang, X.; Wu, Y.; Li, Q.; Yan, J.; Levin, E. Path planning method for USVs based on improved DWA and COLREGs. *Intell. Robot.* **2024**, *4*, 385-405. <http://dx.doi.org/10.20517/ir.2024.23>

Abstract: In the navigation of unmanned surface vehicles (USVs), various types of obstacles may be encountered, which can be categorized into real-time collision avoidance among multiple USVs and obstacle avoidance between USVs and other obstacles. Most existing autonomous obstacle avoidance algorithms do not account for the nonlinear motion characteristics of USVs, often resulting in non-compliance with the International Regulations for Preventing Collisions at Sea (COLREGs) and a tendency to fall into local optima. To address these issues, this paper proposes a path planning algorithm that integrates the dynamic window approach (DWA) considering nonlinear characteristics with COLREGs, making the USV's motion trajectory more applicable to practical engineering scenarios. A kinematic mathematical model is established based on the motion characteristics of USVs, and an evaluation function for the optimal path is constructed using DWA. The fully informed search algorithm (FISA) is employed to select the optimal set of velocities and steering angles from the velocity sampling set, based on different cost calculation methods. The USVs use a laser radar for local obstacle detection, enabling real-time dynamic obstacle avoidance. To address the real-time collision avoidance problem among multiple USVs in open waters, the algorithm filters out COLREGs-compliant avoidance maneuvers during path planning. The correctness and feasibility of the fusion algorithm were verified through comparative simulations. In the simulated environment model, the influence of ocean currents on the USV was introduced, and multiple sets of experiments under different conditions were conducted to compare the motion trajectories, average travel distances, and average travel times of the USV. The simulation results indicate that the USV can perform accurate obstacle avoidance when encountering various types of obstacles. Compared to the traditional DWA algorithm, the proposed approach demonstrates advantages in terms of travel distance and travel time, while still achieving effective obstacle

avoidance.

Keywords: Path planning, improved algorithm, USV, COLREGs

8. Improved DDPG algorithm-based path planning for unmanned surface vehicles

Authors: Menglong Hua, Weixiang Zhou*, Hongying Cheng, Zihao Chen

How to Cite: Hua, M.; Zhou, W.; Cheng, H.; Chen, Z. Improved DDPG algorithm-based path planning for unmanned surface vehicles. *Intell. Robot.* **2024**, *4*, 363-84. <http://dx.doi.org/10.20517/ir.2024.22>

Abstract: As a promising mode of water transportation, unmanned surface vehicles (USVs) are used in various fields owing to their small size, high flexibility, favorable price, and other advantages. Traditional navigation algorithms are affected by various path planning issues. To address the limitations of the traditional deep deterministic policy gradient (DDPG) algorithm, namely slow convergence speed and sparse reward and punishment functions, we proposed an improved DDPG algorithm for USV path planning. First, the principle and workflow of the DDPG deep reinforcement learning (DRL) algorithm are described. Second, the improved method (based on the USVs kinematic model) is proposed, and a continuous state and action space is designed. The reward and punishment function are improved, and the principle of collision avoidance at sea is introduced. Dynamic region restriction is added, distant obstacles in the state space are ignored, and the nearby obstacles are observed to reduce the number of algorithm iterations and save computational resources. The introduction of a multi-intelligence approach combined with a prioritized experience replay mechanism accelerates algorithm convergence, thereby increasing the efficiency and robustness of training. Finally, through a combination of theory and simulation, the DDPG DRL is explored for USV obstacle avoidance and optimal path planning.

Keywords: DDPG deep reinforcement learning, unmanned surface vehicles, obstacle avoidance, path planning

9. Neurodynamics-based formation tracking control of leader-follower nonholonomic multiagent systems

Authors: Xiao-Wen Zhao*, Meng-Ke Li, Qiang Lai, Zhi-Wei Liu

How to Cite: Zhao, X. W.; Li, M. K.; Lai, Q.; Liu, Z. W. Neurodynamics-based formation tracking control of leader-follower nonholonomic multiagent systems. *Intell. Robot.* **2024**, *4*, 339-62. <http://dx.doi.org/10.20517/ir.2024.21>

Abstract: This paper uses a bioinspired neurodynamic (BIN) approach to investigate the formation control problem of leader-follower nonholonomic multiagent systems. In scenarios where not all followers can receive the leader's state, a distributed adaptive estimator is presented to estimate the leader's state. The distributed formation controller, designed using the backstepping technique,

utilizes the estimated leader states and neighboring formation tracking error. To address the issue of impractical velocity jumps, a BIN-based approach is integrated into the backstepping controller. Furthermore, considering the practical applications of nonholonomic multiagent systems, a backstepping controller with a saturation velocity constraint is proposed. Rigorous proofs are provided. Finally, the effectiveness of the presented formation control law is illustrated through numerical simulations.

Keywords: Leader-follower formation control, distributed estimation, nonholonomic multiagent systems, bioinspired neurodynamics, constrained control

10. Intelligent prediction of rail corrugation evolution trend based on self-attention bidirectional TCN and GRU

Authors: Jian-Hua Liu, Wei-Hao Yang, Jing He, Zhong-Mei Wang*, Lin Jia, Chang-Fan Zhang, Wei-Wei Yang

How to Cite: Liu, J. H.; Yang, W. H.; He, J.; Wang, Z. M.; Jia, L.; Zhang, C. F.; Yang, W. W. Intelligent prediction of rail corrugation evolution trend based on self-attention bidirectional TCN and GRU. *Intell. Robot.* **2024**, *4*, 318-38. <http://dx.doi.org/10.20517/ir.2024.20>

Abstract: Analyzing the evolution trend of rail corrugation using signal processing and deep learning is critical for railway safety, as current traditional methods struggle to capture the complex evolution of corrugation. This present study addresses the challenge of accurately capturing this trend, which relies significantly on expert judgment, by proposing an intelligent prediction method based on self-attention (SA), a bidirectional temporal convolutional network (TCN), and a bidirectional gated recurrent unit (GRU). First, multidomain feature extraction and adaptive feature screening were used to obtain the optimal feature set. These features were then combined with principal component analysis (PCA) and the Mahalanobis distance (MD) method to construct a comprehensive health indicator (CHI) that reflects the evolution of rail corrugation. A bidirectional fusion model architecture was employed to capture the temporal correlations between forward and backward information during corrugation evolution, with SA embedded in the model to enhance the focus on key information. The outcome was a rail corrugation trend prediction network that combined a bidirectional TCN, bidirectional GRU, and SA. Subsequently, a multi-strategy improved crested porcupine optimizer (CPO) algorithm was constructed to automatically obtain the optimal network hyperparameters. The proposed method was validated with on-site rail corrugation data, demonstrating superior predictive performance compared to other advanced methods. In summary, the proposed method can accurately predict the evolution trend of rail corrugation, offering a valuable tool for on-site railway maintenance.

Keywords: Mahalanobis distance, rail corrugation, evolution trend prediction, improved crested porcupine optimizer, hybrid time series network

11. An in-vehicle real-time infrared object detection system based on deep learning with resource-constrained hardware

Authors: Tingting Zhuang, Xunru Liang, Bohuan Xue, Xiaoyu Tang*

How to Cite: Zhuang, T.; Liang, X.; Xue, B.; Tang, X. An in-vehicle real-time infrared object detection system based on deep learning with resource-constrained hardware. *Intell. Robot.* **2024**, *4*, 276-92. <http://dx.doi.org/10.20517/ir.2024.18>

Abstract: Advanced driver assistance systems primarily rely on visible images for information. However, in low-visibility weather conditions, such as heavy rain or fog, visible images struggle to capture road conditions accurately. In contrast, infrared (IR) images can overcome this limitation, providing reliable information regardless of external lighting. Addressing this problem, we propose an in-vehicle IR object detection system. We optimize the you only look once (YOLO) v4 object detection algorithm by replacing its original backbone with MobileNetV3, a lightweight feature extraction network, resulting in the MobileNetV3-YOLOv4 model. Furthermore, we replace traditional pre-processing methods with an Image Enhancement Conditional Generative Adversarial Network inversion algorithm to enhance the pre-processing of the input IR images. Finally, we deploy the model on the Jetson Nano, an edge device with constrained hardware resources. Our proposed method achieves an 82.7% mean Average Precision and a frame rate of 55.9 frames per second on the FLIR dataset, surpassing state-of-the-art methods. The experimental results confirm that our approach provides outstanding real-time detection performance while maintaining high precision.

Keywords: Infrared object detection, in-vehicle system, lightweight, limited hardware resources, real-time

12. Parallel implementation for real-time visual SLAM systems based on heterogeneous computing

Authors: Han Liu, Yanchao Dong, Chengbin Hou, Yuhao Liu*, Zhanyi Shu, Sixiong Xu, Tingting Lv

How to Cite: Liu, H.; Dong, Y.; Hou, C.; Liu, Y.; Shu, Z.; Xu, S.; Lv, T. Parallel implementation for real-time visual SLAM systems based on heterogeneous computing. *Intell. Robot.* **2024**, *4*, 256-75. <http://dx.doi.org/10.20517/ir.2024.17>

Abstract: Simultaneous localization and mapping has become rapidly developed and plays an indispensable role in intelligent vehicles. However, many state-of-the-art visual simultaneous localization and mapping (VSLAM) frameworks are very time-consuming both in front-end and back-end, especially for large-scale scenes. Nowadays, the increasingly popular use of graphics processors for general-purpose computing, and the progressively mature high-performance programming theory based on compute unified device architecture (CUDA) have given the possibility for large-scale VSLAM to solve the conflict between limited computing power and excessive computing tasks. The paper proposes a full-flow optimal parallelization scheme based on heterogeneous computing to speed up the time-consuming modules in VSLAM. Firstly, a parallel strategy for feature extraction and matching is designed to reduce the time consumption arising from multiple data transfers between devices. Secondly, a bundle adjustment method based solely on <https://intellrobot.com/>

CUDA is developed. By fully optimizing memory scheduling and task allocation, a large increase in speed is achieved while maintaining accuracy. Besides, CUDA heterogeneous acceleration is fully utilized for tasks such as error computation and linear system construction in the VSLAM back-end to enhance the operation speed. Our proposed method is tested on numerous public datasets on both computer and embedded sides, respectively. A number of qualitative and quantitative experiments are performed to verify its superiority in terms of speed compared to other states-of-the-art.

Keywords: VSLAM, feature extraction and matching, heterogeneous computing, bundle adjustment

13. Importance-driven denial-of-service attack strategy design against remote state estimation in multi-agent intelligent power systems

Authors: Xia Zhao, Guowei Liu*, Lei Li

How to Cite: Zhao, X.; Liu, G.; Li, L. Importance-driven denial-of-service attack strategy design against remote state estimation in multi-agent intelligent power systems. *Intell. Robot.* **2024**, *4*, 244-55. <http://dx.doi.org/10.20517/ir.2024.16>

Abstract: This paper introduces a novel importance-driven denial of service (IDoS) attack strategy aimed at impairing the quality of remote estimators for target agents within multi-agent intelligent power systems. The strategy features two key aspects. Firstly, the IDoS attack strategy concentrates on target agents, enabling attackers to determine the voltage sensitivity of each agent based on limited information. By utilizing these sensitivities, the proposed strategy selectively targets agents with high sensitivity to amplify disruption on the target agent. Secondly, unlike most existing denial of service attack strategies that adhere to predefined attack sequences, IDoS attacks can selectively target important packets on highly sensitive agents, causing further disruption to the target agent. Simulation results on the IEEE 39-Bus system demonstrate that, compared to existing denial of service attack strategies, the proposed IDoS attack strategy significantly diminishes the estimation quality of the target agent, confirming its effectiveness from an attacker's perspective.

Keywords: Multi-agent power systems, remote state estimation, DoS attacks, cyber attacks

14. A novel fatigue driving detection method based on whale optimization and Attention-enhanced GRU

Authors: Zuojin Li*, Minghong Li, Lanyang Shi, Dongyang Li

How to Cite: Li, Z.; Li, M.; Shi, L.; Li, D. A novel fatigue driving detection method based on whale optimization and Attention-enhanced GRU. *Intell. Robot.* **2024**, *4*, 230-43. <http://dx.doi.org/10.20517/ir.2024.15>

Abstract: Fatigue driving has emerged as the predominant causative factor for road traffic safety <https://intellrobot.com/>

accidents. The fatigue driving detection method, derived from laboratory simulation data, faces challenges related to imbalanced data distribution and limited recognition accuracy in practical scenarios. In this study, we introduce a novel approach utilizing a gated recurrent neural network method, employing whale optimization algorithm for fatigue driving identification. Additionally, we incorporate an attention mechanism to enhance identification accuracy. Initially, this study focuses on the driver's operational behavior under authentic vehicular conditions. Subsequently, it employs wavelet energy entropy, scale entropy, and singular entropy analysis to extract the fatigue-related features from the driver's operational behavior. Subsequently, this study adopts the cross-validation recursive feature elimination method to derive the optimal fatigue feature index about operational behavior. To effectively capture long-range dependence relationships, this study employs the gated recurrent unit neural network method. Lastly, an attention mechanism is incorporated in this study to concentrate on pivotal features within the data sequence of driving behavior. It assigns greater weight to crucial information, mitigating information loss caused by the extended temporal sequence. Experimental results obtained from real vehicle data demonstrate that the proposed method achieves an accuracy of 89.84% in third-level fatigue driving detection, with an omission rate of 10.99%. These findings affirm the feasibility of the approach presented in this study.

Keywords: Traffic safety, fatigue driving, operational behavior, whale optimization, neural network

15. Parameter identification of an open-frame underwater vehicle based on numerical simulation and quantum particle swarm optimization

Authors: Mingzhi Chen, Yuan Liu*, Daqi Zhu, Anfeng Shen, Chao Wang, Kaimin Ji

How to Cite: Chen, M.; Liu, Y.; Zhu, D.; Shen, A.; Wang, C.; Ji, K. Parameter identification of an open-frame underwater vehicle based on numerical simulation and quantum particle swarm optimization. *Intell. Robot.* **2024**, *4*, 216-29. <http://dx.doi.org/10.20517/ir.2024.14>

Abstract: Accurate parameter identification of underwater vehicles is of great significance for their controller design and fault diagnosis. Some studies adopt numerical simulation methods to obtain the model parameters of underwater vehicles, but usually only conduct decoupled single-degree-of-freedom steady-state numerical simulations to identify resistance parameters. In this paper, the velocity response is solved by applying a force (or torque) to the underwater vehicle based on the overset grid and Dynamic Fluid-Body Interaction model of STAR-CCM+, solving for the velocity response of an underwater vehicle in all directions in response to propulsive force (or moment) inputs. Based on the data from numerical simulations, a parameter identification method using quantum particle swarm optimization is proposed to simultaneously identify inertia and resistance parameters. By comparing the forward velocity response curves obtained from pool experiments, the identified vehicle model's mean square error of forward velocity is less than 0.20%, which is superior to the steady-state simulation method and particle swarm optimization and genetic algorithm approaches.

Keywords: Underwater vehicle, parameter identification, numerical simulation, quantum particle swarm optimization, dynamic fluid-body interaction
<https://intellrobot.com/>

16. Structural damage identification method based on Swin Transformer and continuous wavelet transform

Authors: Jingzhou Xin, Guangjiong Tao, Qizhi Tang*, Fei Zou, Chenglong Xiang

How to Cite: Xin, J.; Tao, G.; Tang, Q.; Zou, F.; Xiang, C. Structural damage identification method based on Swin Transformer and continuous wavelet transform. *Intell. Robot.* **2024**, *4*, 200-15. <http://dx.doi.org/10.20517/ir.2024.13>

Abstract: The accuracy improvement of deep learning-based damage identification methods has always been pursued. To this end, this study proposes a novel damage identification method using Swin Transformer and continuous wavelet transform (CWT). Specifically, the original structural vibration data is first transferred to a time-frequency diagram by CWT, thereby capturing the characteristic information of structural damage. Secondly, the Swin Transformer is applied to learn the two-dimensional time-frequency diagram layer by layer and extract the damage information, by which the damage identification is achieved. Then, the identification accuracy of the proposed method is analyzed under various sample lengths and different levels of environmental noise to validate the robustness of this approach. Finally, the practicality of this method is verified through laboratory test. The results show the proposed method can effectively recognize the damage and achieve excellent accuracy even under noise interference. Its accuracy reaches 99.6% and 99.0% under single damage and multiple damage scenarios, respectively.

Keywords: Artificial intelligence, deep learning, damage identification, Swin Transformer, continuous wavelet transform

17. Towards environment perception for walking-aid robots: an improved staircase shape feature extraction method

Authors: Xinxing Chen, Yuxuan Wang, Chuheng Chen, Yuquan Leng, Chenglong Fu*

How to Cite: Chen, X.; Wang, Y.; Chen, C.; Leng, Y.; Fu, C. Towards environment perception for walking-aid robots: an improved staircase shape feature extraction method. *Intell. Robot.* **2024**, *4*, 179-95. <http://dx.doi.org/10.20517/ir.2024.11>

Abstract: This paper introduces an innovative staircase shape feature extraction method for walking-aid robots to enhance environmental perception and navigation. We present a robust method for accurate feature extraction of staircases under various conditions, including restricted viewpoints and dynamic movement. Utilizing depth camera-mounted robots, we transform three-dimensional (3D) environmental point cloud into two-dimensional (2D) representations, focusing on identifying both convex and concave corners. Our approach integrates the Random Sample Consensus algorithm with K-Nearest Neighbors (KNN)-augmented Iterative Closest Point (ICP) for efficient point cloud registration. The results show an improvement in trajectory accuracy, with errors within the centimeter range. This work overcomes the limitations of previous approaches and

is of great significance for improving the navigation and safety of walking assistive robots, providing new possibilities for enhancing the autonomy and mobility of individuals with physical disabilities.

Keywords: Walking-aid robots, environment perception, staircase recognition, computer vision, feature extraction

18. A deep learning-based system for accurate detection of anatomical landmarks in colon environment

Authors: Chengwei Ye, Kaiwei Che, Yibing Yao⁴, Nachuan Ma, Ruo Zhang, Yangxin Xu, Jiankun Wang*, Max Q. H. Meng*

How to Cite: Ye, C.; Che, K.; Yao, Y.; Ma, N.; Zhang, R.; Xu, Y.; Wang, J.; Meng, M. Q. H. A deep learning-based system for accurate detection of anatomical landmarks in colon environment. *Intell. Robot.* **2024**, *4*, 164-78. <http://dx.doi.org/10.20517/ir.2024.10>

Abstract: Colonoscopy is a standard imaging tool for examining the lower gastrointestinal tract of patients to capture lesion areas. However, if a lesion area is found during the colonoscopy process, it is difficult to record its location relative to the colon for subsequent therapy or recheck without any reference landmark. Thus, automatic detection of biological anatomical landmarks is highly demanded to improve clinical efficiency. In this article, we propose a novel deep learning-based approach to detect biological anatomical landmarks in colonoscopy videos. First, raw colonoscopy video sequences are pre-processed to reject interference frames. Second, a ResNet-101-based network is used to detect three biological anatomical landmarks separately to obtain the intermediate detection results. Third, to achieve more reliable localization, we propose to post-process the intermediate detection results by identifying the incorrectly predicted frames based on their temporal distribution and reassigning them back to the correct class. Finally, the average detection accuracy reaches 99.75%. Meanwhile, the average intersection over union of 0.91 shows a high degree of similarity between our predicted landmark periods and ground truth. The experimental results demonstrate that our proposed model can accurately detect and localize biological anatomical landmarks from colonoscopy videos.

Keywords: Biological anatomical landmark detection, colonoscopy videos, convolutional neural network (CNN), interference frame rejection, result cleaning

19. Adaptive prescribed performance tracking control for underactuated unmanned surface ships with input quantization

Authors: Jiaming Zhang, Xiang Liu, Xin Wang, Yang Wang*, Yueying Wang*

How to Cite: Zhang, J.; Liu, X.; Wang, X.; Wang, Y.; Wang, Y. Adaptive prescribed performance tracking control for underactuated unmanned surface ships with input quantization. *Intell. Robot.* **2024**, *4*, 146-63. <http://dx.doi.org/10.20517/ir.2024.09>

Abstract: This article investigates the preset performance trajectory tracking control problem of underactuated unmanned surface ships with model uncertainty, unknown external environmental disturbances, and input quantization effects. We consider the non-diagonal damping matrix and mass matrix to satisfy the actual dynamics model of underactuated unmanned surface ships. By adding a hysteresis quantizer, the control method proposed in this article effectively reduces the quantization error. Neural networks are employed to approach the unknown environmental disturbance of underactuated unmanned surface ships. Using the error transformation function, the constrained control problem is transformed into an unconstrained one to ensure the preset performance of tracking errors. This paper verifies the superiority and effectiveness of the proposed control method through Lyapunov stability analysis.

Keywords: Underactuated unmanned surface ships, trajectory tracking control, prescribed performance, neural networks

20. A novel zero-force control framework for post-stroke rehabilitation training based on fuzzy-PID method

Authors: Lina Tong, Decheng Cui, Chen Wang*, Liang Peng*

How to Cite: Tong, L.; Cui, D.; Wang, C.; Peng, L. A novel zero-force control framework for post-stroke rehabilitation training based on fuzzy-PID method. *Intell. Robot.* **2024**, *4*, 125-45. <http://dx.doi.org/10.20517/ir.2024.08>

Abstract: As the number of people with neurological disorders increases, movement rehabilitation becomes progressively important, especially the active rehabilitation training, which has been demonstrated as a promising solution for improving the neural plasticity. In this paper, we developed a 5-degree-of-freedom rehabilitation robot and proposed a zero-force control framework for active rehabilitation training based on the kinematics and dynamics identification. According to the robot motion characteristics, the fuzzy PID algorithm was designed to further improve the flexibility of the robot. Experiments demonstrated that the proposed control method reduced the Root Mean Square Error and Mean Absolute Error evaluation indexes by more than 15% on average and improves the coefficient of determination (R^2) by 4% compared with the traditional PID algorithm. In order to improve the active participation of the post-stroke rehabilitation training, this paper designed an active rehabilitation training scheme based on gamified scenarios, which further enhanced the efficiency of rehabilitation training by means of visual feedback.

Keywords: Upper limb exoskeleton rehabilitation robot, rehabilitation, zero force control, fuzzy control, virtual reality

21. Coordinated energy-efficient walking assistance for paraplegic patients by using the exoskeleton-walker system

Authors: Chen Yang, Xinhao Zhang, Long Zhang, Chaobin Zou*, Zhinan Peng, Rui Huang, Hong Cheng

How to Cite: Yang, C.; Zhang, X.; Zhang, L.; Zou, C.; Peng, Z.; Huang, R.; Cheng, H. Coordinated energy-efficient walking assistance for paraplegic patients by using the exoskeleton-walker system. *Intell. Robot.* **2024**, *4*, 107-24. <http://dx.doi.org/10.20517/ir.2024.07>

Abstract: Overground walking can be achieved for patients with gait impairments by using the lower limb exoskeleton robots. Since it is a challenge to keep balance for patients with insufficient upper body strength, a robotic walker is necessary to assist with the walking balance. However, since the walking pattern varies over time, controlling the robotic walker to follow the walking of the human-exoskeleton system in coordination is a critical issue. Inappropriate control strategy leads to the unnecessary energy cost of the human-exoskeleton-walker (HEW) system and also results in the bad coordination between the human-exoskeleton system and the robotic walker. In this paper, we proposed a Coordinated Energy-Efficient Control (CEEC) approach for the HEW system, which is based on the extremum seeking control algorithm and the coordinated motion planning strategy. First, the extremum seeking control algorithm is used to find the optimal supporting force of the support joint in real time to maximize the energy efficiency of the human-exoskeleton system. Second, the appropriate reference joint angles for wheels of the robotic walker can be generated by the coordinated motion planning strategy, causing the good coordination between the human-exoskeleton system and the robotic walker. The proposed approach has been tested on the HEW simulation model, and the experimental results indicate that the coordinated energy-efficient walking can be achieved with the proposed approach, which is increased by 60.16% compared to the conventional passive robotic walker.

Keywords: Exoskeleton robots, robotic walker, energy efficiency, coordinated motion planning

22. Leveraging active queries in collaborative robotic mission planning

Authors: Cyrille Berger*, Patrick Doherty, Piotr Rudol, Mariusz Wzorek

How to Cite: Berger, C.; Doherty, P.; Rudol, P.; Wzorek, M. Leveraging active queries in collaborative robotic mission planning. *Intell. Robot.* **2024**, *4*, 87-106. <http://dx.doi.org/10.20517/ir.2024.06>

Abstract: This paper focuses on the high-level specification and generation of 3D models for operational environments using the idea of *active queries* as a basis for specifying and generating multi-agent plans for acquiring such models. Assuming an underlying multi-agent system, an operator can specify a request for a particular type of model from a specific region by specifying an active query. This declarative query is then interpreted and executed by collecting already existing data/information in agent systems or, in the active case, by automatically generating high-level mission plans for agents to retrieve and generate parts of the model that do not already exist. The purpose of an active query is to hide the complexity of multi-agent mission plan generation, data transformations, and distributed collection of data/information in underlying multi-agent systems. A description of an active query system, its integration with an existing multi-agent system and validation of the active query system in field robotics experimentation using Unmanned Aerial

Vehicles and simulations are provided.

Keywords: Autonomous robots, active queries, automated planning, situational awareness, 3D reconstruction, unmanned aerial vehicles

23. A small-sample time-series signal augmentation and analysis method for quantitative assessment of bradykinesia in Parkinson's disease

Authors: Zhilin Shu, Peipei Liu, Yuanyuan Cheng, Jinrui Liu, Yuxin Feng, Zhizhong Zhu, Yang Yu, Jianda Han*, Jialing Wu*, Ningbo Yu*

How to Cite: Shu, Z.; Liu, P.; Cheng, Y.; Liu, J.; Feng, Y.; Zhu, Z.; Yu, Y.; Han, J.; Wu, J.; Yu, N. A small-sample time-series signal augmentation and analysis method for quantitative assessment of bradykinesia in Parkinson's disease. *Intell. Robot.* **2024**, *4*, 74-86. <http://dx.doi.org/10.20517/ir.2024.05>

Abstract: Patients with Parkinson's disease (PD) usually have varying degrees of bradykinesia, and the current clinical assessment is mainly based on the Movement Disorder Society Unified PD Rating Scale, which can hardly meet the needs of objectivity and accuracy. Therefore, this paper proposed a small-sample time series classification method (DTW-TapNet) based on dynamic time warping (DTW) data augmentation and attentional prototype network. Firstly, for the problem of small sample sizes of clinical data, a DTW-based data merge method is used to achieve data augmentation. Then, the time series are dimensionally reorganized using random grouping, and convolutional operations are performed to learn features from multivariate time series. Further, attention mechanism and prototype learning are introduced to optimize the distance of the class prototype to which each time series belongs to train a low-dimensional feature representation of the time series, thus reducing the dependency on data volume. Clinical experiments were conducted to collect motion capture data of upper and lower limb movements from 36 patients with PD and eight healthy controls. For the upper limb movement data, the proposed method improved the classification accuracy, weighted precision, and kappa coefficient by 8.89%-15.56%, 9.22%-16.37%, and 0.13-0.23, respectively, compared with support vector machines, long short-term memory, and convolutional prototype network. For the lower limb movement data, the proposed method improved the classification accuracy, weighted precision, and kappa coefficient by 8.16%-20.41%, 10.01%-23.73%, and 0.12-0.28, respectively. The experiments and results show that the proposed method can objectively and accurately assess upper and lower limb bradykinesia in PD.

Keywords: Parkinson's disease, bradykinesia, motion capture, dynamic time warping, attentional prototype network

24. Metropolis criterion pigeon-inspired optimization for multi-UAV swarm controller

Authors: Jinghua Guan*, Hongfei Cheng

How to Cite: Guan, J.; Cheng, H. Metropolis criterion pigeon-inspired optimization for multi-UAV <https://intellrobot.com/>

swarm controller. *Intell. Robot.* **2024**, *4*, 61-73. <http://dx.doi.org/10.20517/ir.2024.04>

Abstract: This paper presents a new multiple unmanned aerial vehicle swarm controller based on Metropolis criterion. This paper presents the design of a controller, utilizing the improved Metropolis criterion pigeon-inspired optimization (IMCPIO) and proportional-integrational-derivative (PID) algorithms, and conducts comparative experiments. Simulation outcomes demonstrate the enhanced performance of the multi-unmanned aerial vehicle formation controller, which is based on IMCPIO, when compared to the basic pigeon-inspired optimization (PIO) algorithm and the genetic algorithm. The IMCPIO algorithm for the energy difference discrimination makes it a faster convergence and more stable effective optimization. Hence, the controller introduced in this study proves to be both practical and resilient.

Keywords: Pigeon-inspired optimization, Metropolis criterion, unmanned aerial vehicle, formation control, proportional-integrational-derivative

25. A collaborative siege method of multiple unmanned vehicles based on reinforcement learning

Authors: Muqing Su, Ruimin Pu, Yin Wang*, Meng Yu

How to Cite: Su, M.; Pu, R.; Wang, Y.; Yu, M. A collaborative siege method of multiple unmanned vehicles based on reinforcement learning. *Intell. Robot.* **2024**, *4*, 39-60. <http://dx.doi.org/10.20517/ir.2024.03>

Abstract: A method based on multi-agent reinforcement learning is proposed to tackle the challenges to capture escaping Target by Unmanned Ground Vehicles (UGVs). Initially, this study introduces environment and motion models tailored for cooperative UGV capture, along with clearly defined success criteria for direct capture. An attention mechanism integrated into the Soft Actor-Critic (SAC) is leveraged, directing focus towards pivotal state features pertinent to the task while effectively managing less relevant aspects. This allows capturing agents to concentrate on the whereabouts and activities of the target agent, thereby enhancing coordination and collaboration during pursuit. This focus on the target agent aids in refining the capture process and ensures precise estimation of value functions. The reduction in superfluous activities and unproductive scenarios amplifies efficiency and robustness. Furthermore, the attention weights dynamically adapt to environmental shifts. To address constrained incentives arising in scenarios with multiple vehicles capturing targets, the study introduces a revamped reward system. It divides the reward function into individual and cooperative components, thereby optimizing both global and localized incentives. By facilitating cooperative collaboration among capturing UGVs, this approach curtails the action space of the target UGV, leading to successful capture outcomes. The proposed technique demonstrates enhanced capture success compared to previous SAC algorithms. Simulation trials and comparisons with alternative learning methodologies validate the effectiveness of the algorithm and the design approach of the reward function.

Keywords: Multi-agent, cooperative capture, soft actor-critic algorithm, attention mechanism, reward function design

26. Pig-ear detection from the thermal infrared image based on improved YOLOv8n

Authors: Hui Han, Xianglong Xue, Qifeng Li*, Hongfeng Gao, Rong Wang, Ruixiang Jiang, Zhiyu Ren, Rui Meng, Mingyu Li, Yuhang Guo, Yu Liu, Weihong Ma*

How to Cite: Han, H.; Xue, X.; Li, Q.; Gao, H.; Wang, R.; Jiang, R.; Ren, Z.; Meng, R.; Li, M.; Guo, Y.; Liu, Y.; Ma, W. Pig-ear detection from the thermal infrared image based on improved YOLOv8n. *Intell. Robot.* **2024**, *4*, 20-38. <http://dx.doi.org/10.20517/ir.2024.02>

Abstract: In the current pig scale breeding process, considering the low accuracy and speed of the infrared thermal camera automatic measurement concerning the pig body surface temperature, this paper proposes an improved algorithm for target detection of the pig ear thermal infrared image based on the YOLOv8n model. The algorithm firstly replaces the standard convolution in the CSPDarknet-53 and neck network with Deformable Convolution v2, so that the convolution kernel can adjust its shape according to the actual situation, thus enhancing the extraction of input features; secondly, the Multi-Head Self-Attention module is integrated into the backbone network, which extends the sensory horizons of the backbone network; finally, the Focal-Efficient Intersection Over Union loss function was introduced into the loss of bounding box regression, which increases the Intersection Over Union loss and gradient of the target and, in turn, improves the accuracy of the bounding box regression. Apart from that, a pig training set, including 3,000 infrared images from 50 different individual pigs, was constructed, trained, and tested. The performance of the proposed algorithm was evaluated by comparing it with the current mainstream target detection algorithms, such as Faster-RCNN, SSD, and YOLO families. The experimental results showed that the improved model achieves 97.0%, 98.1% and 98.5% in terms of Precision, Recall and mean Average Precision, which are 3.3, 0.7 and 4.7 percentage points higher compared to the baseline model. At the same time, the detection speed can reach 131 frames per second, which meets the requirement of real-time detection. The research results show that the improved pig ear detection method based on YOLOv8n proposed in this paper can accurately locate the pig ear in thermal infrared images and provide a reference and basis for the subsequent pig body temperature detection.

Keywords: Thermal infrared image, YOLOv8n, target detection, convolution, MHSA, loss function

27. Recognition of the behaviors of dairy cows by an improved YOLO

Authors: Qiang Bai, Ronghua Gao*, Qifeng Li, Rong Wang, Hongming Zhang

How to Cite: Bai, Q.; Gao, R.; Li, Q.; Wang, R.; Zhang, H. Recognition of the behaviors of dairy cows by an improved YOLO. *Intell. Robot.* **2024**, *4*, 1-19. <http://dx.doi.org/10.20517/ir.2024.01>

Abstract: The physiological well-being of dairy cows is intimately tied to their behavior. Detecting aberrant dairy cows early and reducing financial losses on farms are both possible with real-time and reliable monitoring of their behavior. The behavior data of dairy cows in real environments have dense occlusion and multi-scale issues, which affect the detection results of the model. Therefore, <https://intellrobot.com/>

we focus on both data processing and model construction to improve the results of dairy cow behavior detection. We use a mixed data augmentation method to provide the model with rich cow behavior features. Simultaneously refining the model to optimize the detection outcomes of dairy cow behavior amidst challenging conditions, such as dense occlusion and varying scales. First, a Res2 backbone was constructed to incorporate multi-scale receptive fields and improve the YOLOv3's backbone for the multi-scale feature of dairy cow behaviors. In addition, YOLOv3 detectors were optimized to accurately locate individual dairy cows in different dense environments by combining the global location information of images, and the Global Context Predict Head was designed to enhance the performance of recognizing dairy cow behaviors in crowded surroundings. The dairy cow behavior detection model we built has an accuracy of 90.6%, 91.7%, 80.7%, and 98.5% for the four behaviors of dairy cows standing, lying, walking, and mounting, respectively. The average accuracy of dairy cow detection is 90.4%, which is 1.2% and 12.9% higher than the detection results of YOLOV3, YOLO-tiny and other models respectively. In comparison to YOLOv3, the Average Precision evaluation of the model improves by 2.6% and 1.4% for two similar features of walking and standing behaviors, respectively. The recognition results prove that the model generalizes better for recognizing dairy cow behaviors using behavior videos in various scenes with multi-scale and dense environment features.

Keywords: Dairy cow behaviors, dense environment, YOLOv3, multi-scale, attention module

Contact Us

Email: editorial@intellrobot.com

Twitter: [@OAE_IR](https://twitter.com/OAE_IR)

LinkedIn: [Amber Ren](#)

WeChat: **intellrobot_office**

