RESEARCH

Open Access

Enhancing medical image object detection with collaborative multi-agent deep Q-networks and multi-scale representation

Qinghui Wang¹, Fenglin Liu¹, Ruirui Zou¹, Ying Wang¹, Chenyang Zheng², Zhigiang Tian², Shaoyi Du³ and Wei Zena^{1*}

*Correspondence: zengwei@lyun.edu.cn

¹ School of Physics and Mechanical and Electrical Engineering, Longyan University, Longyan 364012, China ² School of Software Engineering, Xi'an Jiaotong University, Xi'an 710049, China ³ Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an 710049, China

Abstract

Object detection holds a crucial role in medical diagnostics. Tasks like organ segmentation and malignancy diagnosis typically necessitate preliminary localization of corresponding anatomical structures. Precise positioning ensures that only pertinent regions require processing, leading to a potential reduction in computational and storage demands. Conventional image detection approaches necessitate numerous candidate boxes, resulting in redundant computations. Developing techniques capable of accurately detecting medical image objects without reliance on candidate boxes holds substantial practical significance. This paper introduces a 2D method for detecting medical image objects, which leverages multi-agent deep Q-network reinforcement learning and a multi-scale image representation. The method constructs a collaborative environment for multiple agents. These agents individually govern the upper-right corner and lower-left corner positions of the object detection frame, progressively converging toward the actual endpoint through iterative interactions. To expedite the detection process, a multi-scale image representation technique is employed. This method segments the process into three scales. Initially, within the coarse-scale space, the agent approximates the region containing the true endpoint, subsequently executing oscillatory movements. Progressively, it refines its approach within the fine-scale space, advancing toward the genuine endpoint with smaller iterative steps. The detection results demonstrate that collaborative detection among agents yields a 2.45% higher intersection over union compared to non-collaborative detection. Agents exhibit varying step sizes and fields of view in different scale spaces, leading to a reduction in detection time by 0.12 s compared to single-scale comparison. Experimental outcomes demonstrate the superiority of the medical image target detection method proposed in this study over prevailing mainstream detection algorithms.

Keywords: Medical image, Object detection, Deep Q-network, Multi-scale image representation, Multi-agent



© The Author(s) 2023. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http:// creativecommons.org/licenses/by/4.0/.

1 Introduction

In the context of advancing science and technology, the pace of information transmission is rapidly accelerating. In contrast to text, images possess the capability to convey a more vibrant array of information. Image-based data has transcended the confines of text, assuming an indispensable role in the realm of information dissemination-an aspect of utmost significance. Through the analysis of specific images, researchers can extract precise information, thereby furnishing corroborative substantiation for subsequent endeavors.

Image recognition has emerged as a highly prominent research avenue within the realms of computer vision and digital image processing in recent years. This field has captivated numerous scholars, inciting them to engage in its exploration. Concurrently, target detection technology has found widespread applications across various domains of daily life, such as autonomous road navigation and epidemic-induced mask recognition. As a foundational task within computer vision, image recognition underpins a spectrum of other vision-centric objectives, including instance segmentation [1], image interpretation [2], and scene parsing [3]. By efficiently processing image data through computer algorithms, the dual objectives of accuracy and resource conservation are achieved. The ramifications of this endeavor are of substantial practical import, thus elevating image detection technology to the focal point of scientific inquiry. Images, being of diverse typologies, can be broadly categorized into medical images employed for supplementary medical diagnoses, remote sensing images for land resource assessment, and natural images capturing landscapes and individuals. Among these, medical images hold profound relevance to our daily existence. Medical imaging serves as a pivotal tool in contemporary medical research, constituting not only an integral facet of medical progress but also a highly promising avenue within digital imaging technology. Propelled by advancements in medical imaging equipment and the rapid evolution of image processing techniques, the automated computer-based handling of medical images has garnered escalating attention.

The significance of medical images in individuals' lives has been steadily on the rise [4]. Through the utilization of image processing technology, the analysis of medical image data facilitates lesion area detection and instance segmentation within medical images. Medical images assume a paramount role for all, vividly depicting each individual's physiological condition. Regardless of the scale of an issue, whether minor or severe, medical images serve as a conduit for relaying information. The scrutiny of medical image data yields a wealth of crucial insights. Within medical establishments, the detection of medical image targets expedites prompt identification of patient issues, enabling the formulation of tailored treatment strategies, a role of irreplaceable significance in disease diagnosis. On a personal level, the acquisition of medical images during routine physical examinations, coupled with individual medical image reports, empowers individuals to gain foundational insights into their own physical well-being. Within the realm of medical image research, precise localization of target regions augments a multitude of medical image analysis applications. Accurate estimation of target region location and extent enriches subsequent tasks, such as organ segmentation [5, 6], lesion detection [7], and image registration [8], by enabling focus on regions of interest and thereby enhancing algorithm performance. In the context of image segmentation tasks, the pre-detection of

object regions not only elevates segmentation accuracy but also curtails computational demands. Medicine harbors a range of prevalent illnesses; one notable example is prostate cancer, afflicting a considerable portion of male patients and inflicting significant suffering. Radiotherapy stands as the foremost treatment modality for prostate cancer, underscoring the pivotal role of medical image analysis in clinical management. However, the prostate organ's location, dimensions, and contours can exhibit substantial variability across patients. Precisely identifying the prostate organ within medical images poses a formidable challenge, even for seasoned medical practitioners who require time to accurately pinpoint its location. Consequently, the development of an efficient method for object detection in medical images assumes utmost necessity.

Ever since the advent of deep learning, both domestic and international scholars have converged their efforts on this field. Krizhevsky et al. [9] introduced the AlexNet algorithm on the ImageNet dataset in 2012, securing the foremost position across several competition metrics during that period. The progression from AlexNet to ResNet [10], and subsequently to DenseNet [11], underscores an unceasing trajectory of innovation in network architecture and hierarchy. Deep learning has transcended its nascent stages, evolving from rudimentary stacked networks to the inception of diverse high-performance feature extraction modules. Presently, the momentum in deep learning research is intensifying. Domains such as autonomous driving, intelligent recommendations, and unmanned delivery are intricately intertwined with the advancement of deep learning. Remarkable strides have been realized in these domains through the application of deep learning techniques.

Within the realm of image processing, deep learning models exhibit the capability to autonomously extract features, concurrently enhancing the precision of image processing tasks. This advancement significantly mitigates the manual workload associated with feature extraction, yielding notable reductions in costs. With the augmentation of hardware capabilities and the relentless expansion of available data, the landscape of deep learning research in image processing has proliferated expansively. Both domestic and international researchers have contributed distinctive and profound perspectives to the domain of image detection.

In response to the ongoing evolution of reinforcement learning, a synergistic integration of reinforcement learning and deep learning has emerged, leading to significant advancements in diverse domains like gaming, recommendations, and autonomous driving. Capitalizing on the substantial potential showcased by deep reinforcement learning, several scholars have extended its application to the domain of image processing [12– 15]. Within medical image tasks, reinforcement learning is often harnessed to formulate image processing challenges as Markov decision-making scenarios. In the realm of deep reinforcement learning-based target detection, the designated agent's vicinity entails image and contextual information encapsulated within a specified area of interest. This field of view serves as both the agent's input into the network and a pivotal determinant of the ultimate detection accuracy. Over successive interactions between the agent and its environment, the anticipated outcome incrementally materializes. Noteworthy contributions, such as those by Navarro et al. [16], underscore the efficacy of deep reinforcement learning in target detection tasks. These models acquire strategies for object detection through sequential manipulations of bounding boxes, resulting in performance gains. This pursuit of an optimal strategy can be conceptualized as an agent employing a reinforcement learning algorithm to attain maximum rewards through a search process. Within this purview, methods generally yield commendable performance with a relatively small number of iterations. For instance, Stember et al. [17] employed two distinct 3D convolutional neural networks for target detection: one navigates coordinate movements, while the other predicts detection frame dimensions, albeit at a higher computational resource cost. Maicas et al. [18] focused on thymic lesion area detection using deep reinforcement learning, while Kong et al. [19] introduced cooperative communication across Q-networks of different agents, harnessing contextual information for joint detection. Wang et al. [20], on the other hand, proposed an augmented deep neural network for target detection tasks. Grounded in the policy gradient reinforcement learning approach, this method explores various regions within an image, thereby extracting target information during the exploration process.

In the past, most of the detection schemes in medical images were based on machine learning. In recent years, with the popularity of neural networks, the method of object detection through deep learning has become the mainstream. The technology using deep learning requires a large amount of data annotated by experts. A large amount of annotated data is relatively easy to implement in other fields, but in the medical field, data is very precious. On the one hand, the privacy of patients needs to be considered, and the problem of data theft needs to be considered. On the other hand, due to the difficulty of medical image labeling, experienced doctors are required to label, and the training model has high requirements for labeling accuracy. The idea of reinforcement learning can solve some problems. Once the agent interacts with the environment, it will generate a set of data (s, a, s', r) that can be used for training. It can generate a large number of data that can be used on limited medical images. The training data does not require too much labeling information. It only needs to design a reasonable Markov decision-making can be effectively used to solve the problem.

In practical application, the same organ may have varying requirements for the detection frames in terms of size, and an excessive number of candidate frames can also lead to increased consumption of computational resources. In recent years, object detection methods without candidate boxes have emerged [21–25]. In this study, based on the concept of multi-agent deep Q-network (DQN) [26] reinforcement learning, an active exploration method for image detection without candidate frames is proposed. This method is designed to automatically detect target regions in medical images. The model doesn't require candidate frames; instead, it employs two intelligent agents to dynamically detect the coordinates of the lower-left and upper-right corners of the target frame. Through the collaborative efforts of these agents, the target area of the medical image is detected within just a few dozen steps, without the occurrence of redundant candidate boxes. Importantly, this method doesn't necessitate a substantial amount of annotated data during experimentation. The intelligent agents generate a substantial volume of sample data for network updates through continuous interaction with the environment. This approach effectively addresses the limitations of scarce and valuable medical data.

In response to the intricate challenges encountered in current medical image target detection, this paper addresses these issues by delving into the realm of deep reinforcement learning. Specifically, the research delves into the domain of organ detection in medical images and introduces a method tailored for precise medical image target detection. This methodology establishes an intricate environment that facilitates multi-agent collaboration and communication. Within this framework, two agents are assigned distinct roles—managing the upper-right corner and lower-left corner positions of the object detection frame, progressively converging upon the actual endpoint through multiple iterative interactions. In order to expedite the detection process, a multi-scale image representation technique is additionally integrated. This strategy partitions the detection procedure into three discernible scales. In the coarser scale space, the agent initially approximates the region encompassing the actual endpoint. Subsequently, oscillatory movements are employed before gradually zeroing in on the genuine endpoint through more refined steps within the finer scale space.

2 Methods

2.1 2D medical image target detection method based on multi-agent DQN

This section presents an adaptive detection method based on multi-agent collaborative active exploration, which effectively addresses the limitations of preset candidate boxes in deep learning. It has been validated in the detection tasks of 2D medical images. This method ingeniously combines the ideas of deep learning and reinforcement learning, utilizing two agents to detect the coordinates of the lower-left and upper-right vertices of the real bounding box of the target area in 2D medical image. The position and shape of the detection box are delineated by two points on the diagonal, eliminating the need for candidate box design. Through the sharing of convolutional layer parameters, the two agents achieve implicit communication, enabling mutual collaboration and information exchange between agents. This ensures the stability of the environment and the convergence of the network.

The procedural framework of the 2D medical image detection method utilizing the multi-agent reinforcement learning algorithm proposed in this paper is visually depicted in Fig. 1. The two agents share identical convolutional layer network architectures and parameters. The network's input comprises a continuous concatenation of four



Fig. 1 The schematic diagram depicting the image detection process using a multi-agent DQN

consecutive agent state frames, while the output corresponds to the *Q*-values aligned with the established action set of up, down, left, and right movements. By virtue of interactions between the agent and its environment, experiential data represented as s_t^n , a_t^n , r_t^n , s_{t+1}^n (where '*n*' signifies the *n*th agent) is accrued. The current network's *Q*-value prediction, combined with the target network's *Q*-value prediction, is factored in alongside the reward '*r*' for gradient update computation. The pseudocode outlining the DQN algorithm for this multi-agent context is outlined in Fig. 2.

The main idea of this section is deep reinforcement learning. Reinforcement learning truly captured the attention of scholars in 2015 due to the ingenious fusion of deep learning and reinforcement learning [26]. Prior to this, Q-learning was the predominant algorithm in reinforcement learning [27], where the mapping relationship between

Algorithm multi-agent DQN
Input : Current network Q , the target network Q' , the number of experience playback samples m, the number of the updating target network steps C
Output : Parameter θ of the current network Q
function : Multi-agent DQN
1. Initialize the parameter θ of the current network, target network parameter $\theta' \leftarrow \theta$ and
the experience regression pool D
2. for episode=1, M do
3. Initialization state \mathbf{s}_t^1 and \mathbf{s}_t^2 , get the eigenvector $f(\mathbf{s}_t^1) \leftarrow \mathbf{s}_t^1$, $f(\mathbf{s}_t^2) \leftarrow \mathbf{s}_t^2$
4. for $t=1, T$ do
5. for $j=1, 2$ do
6. Input $f(\mathbf{s}_t^J)$ to the current network to get <i>the</i> corresponding action set value Q
7. Use <i>\varepsilon</i> -greedy strategy in the current network to select action a_t^j
8. Execute the current action a_t^j , get the new state s_{t+1}^j and reward $r_t^j f(s_{t+1}^1) \leftarrow s_{t+1}^j$
9. Store $(f(\mathbf{s}_t^j), a_t^j, r_t^j, f(\mathbf{s}_{t+1}^1))$ in the experience regression pool D
10. $\mathbf{S}_t^j = \mathbf{S}_{t+1}^j$
11. Sample m samples from the experience regression pool D and calculate the target network Q' value:
$(r_t^j, ext{ if } s_{t+1}^j \text{ is the final state})$
$\mathbf{y}_{t} = \begin{cases} r_{t}^{j} + \gamma \max a' \mathcal{Q}'(s_{t+1}^{j}, a_{t+1}^{j}, \theta'), & \text{otherwise} \end{cases}$
12. Utilize the mean square error loss function $\frac{1}{m} \sum_{t=1}^{m} (y_t - Q(s_t^j, a_t^j, \theta))^2$ to update
parameter θ of the current network Q m .
13. If t % C = 0, then update the target network $\theta' \leftarrow \theta$
14. endfor
15. endfor
16. end for

Fig. 2 The pseudocode representation of the multi-agent DQN algorithm

states and actions is the primary content recorded in the Q-table. However, due to the limited capacity of the Q-table and the immense volume and complexity of today's data, the Q-table is no longer efficient in accommodating various scenarios. The concept of using neural networks to approximate the Q-table was introduced, and the inception of DQN revolutionized the entire field of reinforcement learning [26]. Addressing the issue that experience data generated by reinforcement learning is serialized and not conducive to network learning, this algorithm introduced the idea of experience replay, enabling independent distribution of data. This algorithm also introduced the notion of a target network corresponding to the current network. The values of the target network are not updated in real time but are periodically copied from the current network, enhancing the stability of the algorithm to some extent.

This paper presents a 2D medical image object detection method grounded in multiagent deep reinforcement learning. Distinguishing itself from single-agent detection strategies, this approach adopts a unique methodology that constructs a cooperative communication framework for multi-agent interaction. This is achieved by permitting the exchange of convolutional layer parameters among all agents. This design is strategically devised to balance environmental stability and algorithmic convergence. Given that each agent is tasked with independently formulating action strategies, the method incorporates two distinct and independent fully connected layers—each agent possessing its own dedicated fully connected layer. The principal functional modules are delineated in "Network structure design based on multi-agent collaborative environment" section.

In reinforcement learning, the design of the Markov decision process is a key component in problem-solving. For the purpose of image detection within the 2D medical environment, we has specifically crafted a Markov decision process, which will be elaborated upon in "Markov decision process design" section. Furthermore, a coarse-to-fine multi-scale image representation method is proposed in this section, which enhances the speed of medical image detection while ensuring the accuracy of 2D medical image target region detection. Through the multi-scale image representation approach, the two agents within the method can roughly detect the area where the target point is located in coarser scale spaces, and then perform precise localization in the finest scale. This approach accelerates the entire detection process to some extent. In "Multi-scale image representation design" section will provide detailed explanations.

2.2 Network structure design based on multi-agent collaborative environment

In single-agent reinforcement learning contexts, the model solely learns from experiences generated by the actions of the individual agent. However, in the realm of multiagent reinforcement learning models, learning derives from experiences originating from the actions of multiple agents within a shared environment. In the framework of the two-agent reinforcement learning (TARL) model introduced in this paper, two agents interact with the environment. For each agent (n = 0 or n = 1) in state s_t^n and taking action a_t^n , the environment elicits a reward signal denoted as r_t^n . The collaborative dynamic between these two agents is depicted in Fig. 3. Notably, when in state 0 and an action is taken, the ensuing state doesn't consistently yield the same outcome due to the influence of Agent 1's interactions with the environment. This manifestation



Fig. 3 The interaction between two agents and the environment in a multi-agent setting



Fig. 4 The implicit communication diagram of two agents

of environmental instability contradicts the Markovian assumption underpinning reinforcement learning scenarios, which rely on the principles of Markov decision processes.

To resolve this dilemma, scholars have advocated for the establishment of inter-agent communication. Such communication entails the sharing of environmental information by any given agent. In the context of multiple agents, communication is often achieved through the implementation of a communication protocol [28]. Moreover, communication among agents can be implicitly accomplished by sharing information within the parameter space [29, 30]. The image detection method articulated in this paper relies on the implicit communication paradigm among agents, and augments the network architecture to facilitate inter-agent communication. The collaborative dynamic among agents centers around the minimization of collective losses, a concept expounded upon in Fig. 4.

The concept of implicit communication between the two agents is graphically represented in Fig. 4. In situations where cooperative communication between agents is absent, each individual agent operates through its own distinct deep Q-network, executing tasks autonomously. Contrasting this, the method designed in this paper forges a novel approach by deploying two collaborating agents that share a convolutional neural network. Inter-agent information interchange is accomplished through the mutual sharing of convolutional layer parameters. The network architecture, rooted in ResNet 18 [10], lays the foundation for an enhanced network model, as depicted in Fig. 5. This paper constitutes 17 convolutional layers and is supplemented by 2 fully connected layers. While the convolutional layer parameters are shared between the two agents, each



Fig. 5 The ResNet 18 network structure diagram

agent possesses an independent fully connected layer. This design choice is attributed to the distinct role of the fully connected layer in shaping the final action policy for each agent.

The convolutional layer's role in the network is twofold-it facilitates the learning of underlying features while accommodating the concurrent processing of two inputs. In contrast, the fully connected layer is adept at capturing advanced features, with the retention of target point information. A pivotal outcome of sharing parameters among convolutional layers is the reduction in the total number of parameters, streamlining computations and preserving crucial contextual information. This parameter sharing strategy also enables agents to indirectly communicate their experiences within the parameter space.

2.3 Markov decision process design

Reinforcement learning constitutes a dynamic process where an agent assumes the role of a decision-maker, continuously interacting with a specific environment to make sequential decisions. Prior to constructing the detection method, it is essential to formulate a Markov decision process tailored to the task scenario. The design of this Markov decision process four key elements: the state set, action set, reward function, and conditions for detection termination.

- (1) State Set *S*: The framework assumes the agent operates within a 2D medical image environment. Within the state set *S*, each state *s* corresponds to pixels encompassed by the detection frame $b = [b_{x1}, b_{y1}, b_{x2}, b_{y2}]$. This involves the coordinates of the upper-right corner (b_{x1}, b_{y1}) and lower-left corner (b_{x2}, b_{y2}) of the detection frame. To ensure stability within the agent's detection process, the input state for the network is crafted by concatenating the current state with its three nearest states across channels.
- (2) Action Set A: The section outlines a collection of discrete actions $A = \{a_x^+, a_x^-, a_y^+, a_y^-\} \in \mathbb{R}^4$. In this context, each agent can undertake positive and negative movements along the *x* or *y* axis, manifesting as four actions: right, left, up, and down. These four actions align with the outputs of the previously mentioned fully connected layer of the network. The network designed within this paper yields a four-dimensional output, corresponding to the four actions outlined above. As illustrated in Fig. 6, the blue dot signifies the agent's position, while the



Fig. 6 An example of actions taken by agents in a multi-agent setting

yellow border delineates the agent's state. By executing these four distinct actions, the agent attains the capability to navigate every position within the environment.

In the proposed multi-agent DQN reinforcement learning approach for medical image object detection, the actions of each agent corresponds to the movement of one of the four possible directions on the two corner points, influencing the size and position of the detection frame and facilitating precise localization of the object.

Here's a more detailed explanation of how each agent's action affects the detection frame:

(a) Up and down movements:

- Up movement: When an agent chooses to move up, it shifts the detection frame's upper edge upwards, effectively reducing the frame's height. This action can be seen as focusing on the region higher in the image.
- Down movement: Conversely, when an agent moves down, it shifts the detection frame's lower edge downward, increasing the frame's height. This action concentrates on the region lower in the image.
- (b) Left and right movements:
 - Left movement: Moving left shifts the left edge of the detection frame to the left, decreasing its width. This action can be thought of as focusing on the region to the left of the image.
 - · Right movement: Moving right shifts the right edge of the detection frame to the right, increasing its width. This action concentrates on the region to the right of the image.

- (c) Combination of movements:
- The agents can also perform combinations of these movements in multiple steps. For example, an agent can simultaneously move up and left, leading to a diagonal shift in the detection frame. This allows for more fine-grained adjustments of the frame's size and position to accurately match the object's location within the medical image. These movements provide a flexible mechanism for dynamically sizing and positioning the detection frame, ultimately leading to precise and efficient object localization. The collaborative nature of the multi-agent system ensures that both agents work together to gradually refine the detection frame's position and dimensions until it accurately encompasses the object. This dynamic approach eliminates the need for predefined candidate boxes, resulting in a more efficient and precise detection process.
- (d) Reward Functions: The formulation of reward functions is a pivotal aspect aimed at maximizing cumulative rewards across sequential interactions. Reward functions play a pivotal role in motivating agents to perform specific actions as directed by the policy, thereby moving in desired directions. Given that this paper's method revolves around detecting target areas based on key points, a comprehensive approach to designing reward functions is adopted. To facilitate effective medical image target detection, the design considers multiple aspects including the Euclidean distance between the current agent's represented detection point and the target frame. Additionally, variations in the cross-merge ratio are factored in. Three distinct reward functions are introduced, as depicted in Eq. (1).

$$R_1 = O_{t+1} - O_t$$

$$R_2 = IoU_{t+1} - IoU_t$$

$$R_3 = BO_{t+1} - BO_t$$
(1)

At time step 't', the Euclidean distance between the current agent's represented detection point and the target point is denoted as ' O_t '. The detection points represented by both agents form a detection box, with the intersection ratio of the detection frame and the target frame being 'IoU_t'. The Euclidean distance between the coordinates of the center point of the detection frame and the center point of the detection frame and the center point of the target frame is designated as 'BO_t'.

(e) Detection termination conditions: The proposed method employs distinct termination conditions for detection during training and testing phases. In the training process, if the distance between the detection points represented by the two agents and the target point equals or falls below a predefined threshold, the ongoing detection sequence concludes, resulting in a successful detection. Alternatively, if the maximum number of steps has been reached and yet the Euclidean distance value between the agent's detection point and the target point remains above the threshold, the detection is considered unsuccessful. During the testing process, oscillation detection serves as the termination criterion. If the agent's represented coordinate point appears in the smallest scale space more than three times, it indicates oscillatory behavior during the agent's detection process. Subsequently, the detection is halted, with the current oscillation point serving as the final detection outcome.

2.4 Multi-scale image representation design

Inspired by relevant work [12], the method of multi-scale image representation is used in this paper to further improve the speed of image detection while ensuring the detection accuracy of medical image target regions.

The concept of multi-scale image representation encompasses two aspects. Firstly, the agent's step size operates across multiple scales. Within the design, when the two agents are in the coarsest scale space (M scale space), their movement step is set to 9 voxels. In the M-1 scale space, the agent's movement step becomes 3 voxels, and in the smallest scale space (M-2 scale space), the agent's movement step is reduced to 1 voxel. Secondly, the agent's field of view extends across multiple scales. This field of view, centered on the agent's current position, is represented by pixels within a specified size area surrounding the agent. The formula for visual field transformation across multiple scale spaces is given by Eq. (2):

$$x_{\min} = x_{loc} - \frac{w * x_s}{2}, \ x_{\max} = x_{loc} + \frac{w * x_s}{2}$$

$$y_{\min} = y_{loc} - \frac{h * y_s}{2}, \ y_{\max} = y_{loc} + \frac{h * y_s}{2}$$
(2)

where (x_{\min}, y_{\min}) and (x_{\max}, y_{\max}) represent the coordinates of the two endpoints on the diagonal of the agent's field of view frame, (x_{loc}, y_{loc}) represents the current position coordinates of the agent, w and h are hyperparameters, both set to 225, and x_s and y_s represent the scales along the *x* and *y* axes. In the three-scale spaces, x_s and y_s maintain consistent values: 3, 2, and 1.

As depicted in Fig. 7, the agent 0 (Agent 0) is the investigation subject. The yellow border signifies the agent's field of view across different scale spaces, while the red dot



Fig. 7 Multi-scale oscillation detection diagram

denotes the target point. The blue dot represents the agent's current position (current detection point). The color of the "Error" label in the lower-left corner serves as an indicator-green implies the agent moved closer to the target point after the last action, while red signifies a greater distance. If the agent occupies a certain position in a particular scale space three times, the model concludes that oscillation transpired during the agent's detection process.

Refer to Fig. 8 for visual representations, subgraphs (a) and (b) depict the detection process of the two agents across different scale spaces. In the same scale space, the left and right panels visualize the detection of agent 0 and agent 1, respectively.

Distinct oscillation detection models are tailored for each scale space level, catering to the diverse demands of different scales. The detection process initiates from the coarsest scale level, denoted as M. Within this level, the agent's field of view is expansive,



(a) Scale space=M



(b) Scale space=M-1



 $(c) \ Scale \ space=M-2$ Fig. 8 Illustration of collaborative detection by two intelligent agents

facilitating the acquisition of extensive global context information that ensures efficient navigation. Concurrently, the agent's movement step is substantial. Upon the agent's arrival at the target point or when oscillation manifests within the vicinity of the target point, it is deduced that the agent has achieved convergence within the current scale space. As a result, the scale space transitions to M-1. Continuing the detection process, exploration recommences from the convergence point within the preceding scale space. Simultaneously, the agent's field of view is reduced along with a decrease in the step size. This adjustment prevents scenarios where a large step size might cause the agent to overshoot the target point, enabling more accurate target point localization. This iterative process repeats across subsequent scales. Should oscillations become apparent even in the finest scale, the agent's detection endeavor culminates.

3 Experiments and analysis

3.1 Dataset and evaluation criteria

The experimental phase encompassed a quantitative assessment of the proposed detection method, conducted across two distinct 2D medical image datasets: the amalgamated collection of prostate medical images [31] and the REFUGE fundus dataset [32]. A series of experiments were performed, including ablation and comparison analyses. The hybrid dataset employed for prostate medical imaging consisted of 3795 slices, amalgamated from three distinct prostate MR datasets. Notably, the training dataset comprised 3371 slices originating from the ISBI2013, PROMISE12, and Emory datasets. For testing purposes, 424 slices from the PROMISE12 test dataset were employed. All images were meticulously annotated by proficient radiologists, rendering this test set a widely recognized benchmark extensively employed for evaluating diverse algorithms within the medical community. The REFUGE fundus dataset, akin to PROMISE, was sourced from an open-access competition dataset, further underpinning its established prominence.

The results presented in "Experiment and analysis" section are the average outcomes derived from multiple experiment repetitions. The experiments were conducted with diverse training and testing image sets to ensure the robustness of the proposed method. Each experiment iteration involved training the model on a distinct dataset and evaluating its performance on another dataset. This process was repeated more than ten times to ensure statistical robustness and account for potential variations in the datasets. The choice of the number of repetitions was determined based on considerations of model complexity and the desire for reliable and representative results.

As for the choice of using only the prostate images from the PROMISE12 dataset as the testing set, there are specific reasons for this decision:

- (1) Focus on specific task: We have designed our method with a particular focus on prostate image detection. By concentrating the testing set on prostate images, we can better evaluate the method's effectiveness for this specific task.
- (2) Comparison with existing approaches: It is common in research to compare a new method's performance against existing approaches. By using the same testing set that previous methods used, we can directly compare their results to those of others, providing a basis for assessing the method's competitiveness.

- (3) Availability of ground truth: The PROMISE12 dataset may provide high-quality ground truth annotations for prostate images, which are crucial for training and evaluating object detection methods. If this dataset offers comprehensive annotations for the task at hand, it becomes a valuable resource for testing the proposed method.
- (4) Standardization: Using a standardized testing dataset can make it easier for other researchers to replicate the experiments and compare their methods with the proposed one, contributing to the reproducibility of research.

It is not uncommon for research papers to use specific datasets for evaluation, as long as the choice is well-justified and aligns with the research objectives.

To assess the experimental outcomes, this study employs three predominant evaluation metrics commonly applied in medical image target detection [33, 34]: the IoU ratio, wall distance (WD), and centroid distance (CD) between the detection frame and the target frame.

(1) The formula for calculating the IoU metric is shown as Equation (3):

$$IoU_{A,B} = \frac{S_A \cap S_B}{S_A \cup S_B}$$
(3)

where S_A denotes the area of the target frame, S_B signifies the area of the detection frame, $S_A \cap S_B$ corresponds to the intersection area of the target frame and the detection frame, and $S_A \cup S_B$ signifies the union area of the target frame and the detection frame.

(2) The average WD signifies the mean absolute distance between the four sides of the detection frame and the corresponding sides of the target frame. This metric offers a direct assessment of the alignment between the detection frame and the target frame in the approach. The calculation formula is presented in Equation (4):

$$D_{\text{wall dist}} = \frac{D_{up} + D_{\text{down}} + D_{\text{left}} + D_{\text{right}}}{4} \tag{4}$$

where D_{up} , D_{down} , D_{pleft} , and D_{right} symbolize the absolute distances of the four pairs of bounding boxes: the upper, lower, left, and right sides of the detection frame and the target frame, respectively.

(3) CD denotes the centroid distance between the detection frame and the target frame, which is the Euclidean distance between the centroid coordinates of the two frames, quantifying the disparity in the central region. The calculation formula is illustrated in Eq. (5):

$$D_{\text{centroid dist}} = \sqrt{\left(\frac{x_1 + x_2}{2} - \frac{x_1^{'} + x_2^{'}}{2}\right)^2 + \left(\frac{y_1 + y_2}{2} - \frac{y_1^{'} + y_2^{'}}{2}\right)^2}$$
(5)

where (x_1, y_1) represents the coordinates of the upper-right corner of the target frame, while (x_2, y_2) signifies the coordinates of the lower-left corner of the target frame. Similarly, (x'_1, y'_1) refers to the upper-right corner coordinates of the

detection frame, and (x'_2, y'_2) represents the coordinates of the lower-left corner of the detection frame.

3.2 Experimental parameters

All code implementations of the proposed detection method in this paper have been successfully executed in Python, and all experiments were conducted using a GPU with 11 GB of memory. In the initial 10 epochs, the agent's exploration rate decreases from 1 to 0.1. Additionally, the reward discount factor γ is set to 0.9. During the training process, for each 2D medical image, in order to augment the sample size, the agent can initiate from multiple starting points. Furthermore, to ensure the initial point remains within the image boundaries, the positional coordinates for each initial point can only be selected within a fixed area at the center of the image. The network outputs consist of four values, corresponding to the four actions of the agent. The proposed method employs the ε -Greedy strategy, wherein if the generated random number surpasses the exploration rate ε , the action corresponding to the maximum *Q*-value is taken; otherwise, a random action is chosen. Once the agent enters an oscillatory state, the scale space diminishes. If the minimum scale space is reached, the detection process terminates.

3.3 Result visualization

Throughout the training process, the experiment's images are evenly partitioned into nine sections, with the initial detection points for both agents being selected at the centroid positions of the upper, lower, left, right, and middle sections of the 2D medical image. This is illustrated in Fig. 9, subfigure (a). This approach not only enhances the training dataset but also mitigates the occurrence of agent initial points situated near the image's boundaries, thus reducing the likelihood of moving out of the image. During testing, as depicted in Fig. 9, subfigure (b), the centroid coordinates of the entire 2D medical image serve as the sole initial point for the agent. The visualization of detection results across distinct datasets is presented in Fig. 10.



(a) Training initial point selection illustration
 (b) Test initial point selection illustration
 Fig. 9 Agent initial point selection graph: a Training initial point selection illustration, and b Test initial point selection illustration



(a) REFUGE fundus dataset
 (b) Prostate dataset
 Fig. 10 Visualization of detection results on different datasets: a REFUGE fundus dataset, and b Prostate dataset

In Fig. 10, the visual representations are demonstrated for both the prostate dataset and the fundus dataset, respectively, where the detected bounding boxes are in close proximity to the target bounding boxes. The green bounding boxes depict the ground truth bounding boxes, while the red bounding boxes indicate the detected bounding boxes. For subfigure (a), showcasing the REFUGE fundus dataset, the smaller optic disk area is magnified and displayed. The preliminary assessment of the detection results through visual representation suggests that the multi-agent-based 2D medical image target detection approach utilizing the DQN reinforcement learning algorithm holds practical significance and feasibility. The method proposed in this paper will receive further validation through ablation experiments and comparative analyses leveraging experimental data.

3.4 Ablation experiment

The subsequent ablation experiments were conducted exclusively on the mixed prostate dataset. Scale ablation experiments were carried out, comparing across various scales. In the single-scale experiment, setting the agent's step size too large could result in the agent crossing the target point during movement. Hence, in this experiment, a step size of one pixel was employed, with the entire image serving as the agent's field of view. The comprehensive explanation of the multi-scale design has been previously presented and will not be reiterated here. Within the specific experimental process, ablation experiments were also performed across a range of multi-scale scenarios, encompassing two-scale exploration, three-scale exploration, and four-scale exploration.

The comparative outcomes of both multi-scale and single-scale experiments are outlined in Table 1. In the single-scale experiment, the complete 2D medical images were adopted as the agent's field of view. This approach, however, resulted in overly sparse input information for the network, thereby adversely impacting the detection performance, with an IoU of only 63.15%. Furthermore, the agent's step size remained fixed at one pixel, leading to a relatively slower detection speed. In the two-scale experiment, the rapid alteration in the scale space rendered the two-scale exploration methodology

Scale	Average IoU (%)↑	Wall distance (mm) \downarrow	Center of mass distance (mm)↓	Time (s)↓
Single-scale	63.15	11.70±12.35	9.63±8.99	0.50
Two-scale	75.64	6.87±6.53	6.45±6.32	0.41
Three-scale	80.70	$\textbf{4.39} \pm \textbf{3.50}$	$\textbf{4.36} \pm \textbf{4.15}$	0.38
Four-scale	80.15	4.94±4.76	4.67±5.07	0.43

Table 1	Single	-scale	and	multi	-scale	ablatior	n experime	ents

The bold is used to highlight the best results in the comparisons

Table 2 Multi-agent communication ablation experiments

	Average IoU (%)↑	Wall distance (mm)↓	Center of mass distance (mm)↓
Multi-agent	78.25	5.70 ± 6.15	5.16 ± 4.78
Multi-agent + implicit com- munication	80.70	$\textbf{4.39} \pm \textbf{3.50}$	$\textbf{4.36} \pm \textbf{4.15}$

The bold is used to highlight the best results in the comparisons

ineffective in capitalizing on contextual information, thus yielding suboptimal detection outcomes. Nevertheless, this approach contributed to a certain degree of acceleration in the detection process. The performance of the four-scale exploration approach mirrored that of the three-scale counterpart, although detection time started to increase due to the impact of the shock detection strategy applied in the experiment. As the number of scale spaces gradually increased, the agent's need to undertake oscillation detection in each scale space led to a rise in detection time.

Remarkably, medical image detection demanded merely 0.38 s. When juxtaposed with several alternative scale image representation techniques, the three-scale image representation strategy not only ensured precision in medical image object detection but also enhanced the detection speed. Additionally, this study conducted a multi-agent experiment concerning prostate region detection, devoid of collaborative communication. A comparison was drawn against multi-agent cooperative detection of the prostate region, facilitated by convolutional layer parameter sharing. In the multi-agent non-cooperative experiments, a unique depth was constructed for each of the two agents in the Q-network. These agents, respectively, pinpointed the upper-right and lower-left corners of the target frame. Notably, these agents did not engage in cooperation or information exchange throughout the detection process. Conversely, the experiment involving multiagent cooperative communication detection involved two agents collaboratively detecting the upper-right and lower-left corners of the target frame. In tandem with fulfilling their individual detection tasks, these agents exchanged information with each other, leading to more refined detection outcomes. The parameter sharing approach concurrently curtailed computational complexity. For specific experimental outcomes, please refer to Table 2.

The results of the experimental comparison between multi-agent cooperation and non-cooperation are presented in Table 4. The experimental findings underscored the efficacy of multi-agent collaboration in enhancing the accuracy of prostate region

Network backbone	Average IoU (%)↑	Wall distance (mm) \downarrow	Center of mass distance (mm)↓
DenseNet [11]	80.46	4.52 ± 3.71	4.44 ± 4.55
ResNet 18 [10]	80.70	$\textbf{4.39} \pm \textbf{3.50}$	$\textbf{4.36} \pm \textbf{4.15}$

 Table 3
 Multi-agent communication ablation experiments

The bold is used to highlight the best results in the comparisons

Methods Average IoU (%) Wall distance (mm)↓ Center of mass distance (mm) ↓ Faster R-CNN [35] 80.23 4.63 ± 3.98 4.43 ± 4.86 SSD300 [36] 70 12.24 ± 29.20 14.02 ± 40.86 YOLOv3 [37] 78.34 7.70 ± 21.50 8.95 ± 31.11 DETR [38] 80.36 4.53 ± 4.79 4.41 ± 4.92 DDPG [39] 65.75 9.60 ± 6.90 9.26 ± 6.11 YOLOv4 [40] 80.04 4.75 ± 4.12 4.46 ± 5.58 Our method 80.70 $\textbf{4.39} \pm \textbf{3.50}$ 4.36 ± 4.15

Table 4	Detection	results of	prostate	dataset	using	different	methods

The bold is used to highlight the best results in the comparisons

detection. Moreover, as introduced earlier, both DenseNet and ResNet networks were subjected to testing and exploration in the course of this experimentation. The ultimate outcomes of these explorations are encapsulated in Table 3.

In summary, the comparison of the abovementioned ablation experiments allows us to deduce that the novel strategy embraced by the method advocated in this paper has the potential to enhance performance in intricate medical settings. This, in turn, can bolster the detection accuracy of 2D medical images.

3.5 Comparative experiments with existing methods

In this study, the proposed method is compared with several other mainstream image detection methods, including Faster R-CNN [35], SSD300 [36], YOLOv3 [37], DETR [38], DDPG [39], and YOLOv4 [40]. To assess the detection accuracy of the method proposed in this paper, not only is the average IoU between the ground truth frame marked by experts and the detection frame calculated, but also the wall distance and centroid distance between the detection frame and the target frame are computed. Metrics for wall distance and centroid distance are presented as mean \pm standard deviation. Table 4 presents the average IoU, wall distance, and centroid distance metrics of different methods on the prostate dataset, showing that SSD and YOLOv3 have relatively poor detection performance. The experimental results demonstrate that the proposed method performs exceptionally well across multiple indicators, achieving the highest average IoU and the smallest error in wall distance and centroid distance.

Analysis of the data in Table 4 leads to the conclusion that the multi-agent DQN reinforcement learning detection method proposed in this paper excels in complex medical image environments, exhibiting superior results in terms of intersection ratio, wall distance, and centroid distance indicators. In comparison with other methods, DDPG demonstrates that its deterministic strategy algorithm does not perform as well as the DQN algorithm in discrete action scenarios, albeit with fewer discrete anomalies. Due to the limited availability of open-source resources for reinforcement learning in image processing, most comparison experiments in this paper rely on deep learning. Notably, SSD's detection performance on the dataset is subpar, whereas DETR and Faster R-CNN algorithms achieve relatively favorable results. The proposed method achieves an intersection ratio of 80.07%, outperforming mainstream target detection algorithms particularly on the prostate dataset. The low standard deviation values demonstrate the method's robustness. To visualize detection errors more intuitively, box plots depict wall distance and centroid distance data. Figure 11 displays a boxplot representing prediction deviations, with discrete points indicating abnormal detection values. Notably, the proposed method exhibits fewer abnormal data points, aligning with expectations and highlighting the suitability of the multi-agent reinforcement learning approach for intricate medical image environments and 2D medical image target detection, given its high robustness.

To further validate the proposed method's robustness, the REFUGE dataset [32] is used for comparative experiments, focusing on the optic disk area in the fundus dataset. The optic disk region's detection employs the exact parameters used for the prostate dataset, and training is conducted on the same server. As detailed in Table 5, the detection results for the optic disk area in the REFUGE fundus dataset underscore the superior performance of the proposed method. It achieves optimal metrics across all indicators, with an average intersection ratio of 79.94%. The wall distance and centroid distance indicators are also better than those of the DETR algorithm, exhibiting a lower margin of approximately 0.3 mm. The experimental findings underscore the high applicability of the proposed image detection method on 2D medical image datasets.

The primary focus of our study lies in the introduction and evaluation of a novel collaborative multi-agent DQN model for medical image object detection. While the



Fig. 11 Boxplot of the detection results of the prostate dataset

Methods	Average loU (%)↑	Wall distance (mm)↓	Center of mass distance (mm)↓
Faster R-CNN [35]	78.63	6.16	7.74
YOLOv3 [37]	75.26	9.23	9.67
DETR [38]	79.35	5.68	6.29
DDPG [39]	64.47	10.89	10.76
YOLOv4 [40]	76.89	8.59	8.20
Our method	79.94	5.37	6.59

Table 5 Detection results of REFUGE fundus dataset using different method
--

The bold is used to highlight the best results in the comparisons

integration of various deep reinforcement learning (DRL) methods could indeed contribute to a broader comparative analysis, several considerations guided our decision to specifically exclude certain DRL methods, including the work by Ghesu et al. [12].

- (1) Methodological coherence: Our study aimed for methodological coherence and depth rather than breadth. Including multiple DRL methods could introduce significant variations in architectures, training strategies, and hyperparameters, complicating the interpretation of results. By concentrating on a single DRL model, we sought to provide a clear and detailed understanding of its performance in the context of medical image detection.
- (2) Unique contribution: The proposed multi-agent DQN model represents a unique contribution to the field, emphasizing collaborative detection in medical images. Introducing too many comparative methods might dilute the focus on our novel approach. We believe that an in-depth analysis of the proposed model's performance, both quantitatively and qualitatively, provides valuable insights for the targeted application.
- (3) Resource limitations: Given the constraints on resources, including space limitations in the manuscript and computational resources required for extensive experiments, a strategic choice was made to ensure a comprehensive yet manageable evaluation. This allowed us to delve deeply into the proposed model's capabilities and limitations.
- (4) Experimental stability: Maintaining experimental stability is crucial for drawing meaningful conclusions. Focusing on a specific DRL model, along with YOLOv3 and YOLOv4, provided a stable and controlled environment for evaluation.

In essence, our rationale for not including other works based on DRL methods was to prioritize depth, clarity, and coherence in the evaluation of the proposed multi-agent DQN model in the specific domain of medical image object detection.

While our proposed method exhibits great potential, it does come with inherent limitations. These include computational demands, sensitivity to hyperparameters, constraints related to data quality and diversity, as well as challenges concerning overfitting and interpretability. Efforts are required to address these limitations in future research. Optimization strategies for computational efficiency, adaptive algorithms, enhanced data diversity and quality, applicability expansion across various medical tasks, measures against overfitting, and improved model interpretability will be explored. These endeavors aim to refine and advance our innovative approach, making it a valuable asset for practical medical applications.

4 Conclusion

This paper introduces a novel image detection methodology that capitalizes on the multi-agent DQN reinforcement learning algorithm and multi-scale image representation. This approach operates on the premise of target detection sans candidate bounding boxes, achieving this by pinpointing two pivotal locations. However, it necessitates cooperative communication between agents, each representing one of these key locations. This technique introduces several innovations and contributions. Firstly, it advocates for the utilization of two agents to collaboratively identify two key points along the diagonal of the target region within 2D medical images. These points subsequently define the detection frame, eliminating the need to regress the entire frame. Consequently, this obviates redundant candidate frame computations. Secondly, by facilitating parameter sharing among agents' convolutional layers, an interactive communication environment is cultivated, fostering multi-agent collaboration. This cooperative setting enables agents to mutually benefit from shared experiences, yielding superior detection outcomes. Notably, cooperative detection surpasses non-cooperative detection by 2.45% based on the results. Thirdly, a progressive multi-scale image representation technique is introduced. When an agent encounters oscillations in a coarser scale space, the convergence criteria are applied within that scale space before transitioning to the next phase of training. Agents exhibit varying step sizes and field of view across different scale spaces, yielding a reduction of 0.12 s in detection time compared to single-scale approaches. Lastly, the proposed method attains the highest intersection-over-union ratio, alongside the lowest wall distance and centroid distance results in the prostate hybrid dataset and fundus dataset. The method, while commendable, bears prospects for enhancement. Enhancements can be pursued through network structure refinements, incorporating the convolutional layer parameter sharing technique to facilitate implicit communication between agents. In future endeavors, the fully connected layer might benefit from an average parameter approach to foster communication, while the last fully connected layer could retain a degree of independence.

Abbreviations

- DQN Deep Q-network
- IoU Intersection over union
- TARL Two-agent reinforcement learning
- WD Wall distance
- CD Centroid distance

Acknowledgements Not applicable.

Author Contributions

WZ performed conceptualization; QW provided methodology; FL and YW provided software; RZ and CZ performed validation; QW and CZ performed writing; ZT performed writing—review and editing; SD performed project administration; SD contributed to funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by the Natural Science Foundation of Fujian Province (Grant Nos. 2021J011086, 2022J011146, 2023J01964, 2023J01965, 2023J01966), by the Natural Science Basic Research Program of Shaanxi (Grant No. 2021JM-020), by the Fujian Province Chinese Academy of Sciences STS Program Supporting Project (Grant

No. 2023T3084), by the Guidance Project of the Science and Technology Department of Fujian Province (Grand No. 2023H0017), by the Xinluo District Industry-University-Research Science and Technology Joint Innovation Project (Grand Nos. 2022XLXYZ002,2022XLXYZ004), by the Key Project of Shaanxi Province (Grand No. 2018ZDCXL-GY-06-07), and by the Qimai Science and Technology Innovation Project of Wuping Country.

Availability of data and materials

The datasets used and/or analyzed during in current study are available from the corresponding author on reasonable requests.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 14 September 2023 Accepted: 10 December 2023 Published online: 21 December 2023

References

- K. He, G. Gkioxari, P. Dolláir, R. Girshick, Mask, R-CNN, in Proceedings of the IEEE International Conference on Computer Vision (2017), pp. 2961–2969
- P. Anderson, X. He, C. Buehler, D. Teney, M. Johnson, S. Gould, L. Zhang, Bottom-up and top-down attention for image captioning and visual question answering, in *Proceedings of the IEEE International Conference on Computer Vision* (2018), pp. 6077–6086
- 3. J. Yang, J. Lu, S. Lee, D. Batra, D. Parikh, Graph R-CNN for scene graph generation, in *Proceedings of the European* Conference on Computer Vision (2018), pp. 670–685
- M.A. Abdou, Literature review: efficient deep neural networks techniques for medical image analysis. Neural Comput. Appl. 34(8), 5791–5812 (2022)
- 5. P.N. Samarakoon, E. Promayon, C. Fouard, Light random regression forests for automatic multi-organ localization in CT images, in *IEEE 14th International Symposium on Biomedical Imaging* (2017), pp. 371–374
- S. Liang, K.H. Thung, D. Nie, Y. Zhang, D. Shen, Multi-view spatial aggregation framework for joint localization and segmentation of organs at risk in head and neck CT images. IEEE Trans. Med. Imaging 39(9), 2794–2805 (2020)
- 7. T. Wang, G. Liao, L. Chen, Y. Zhuang, S. Zhou, Q. Yuan, M. Zhang, Intelligent diagnosis of multiple peripheral retinal lesions in ultra-widefield fundus images based on deep learning. Ophthalmol. Ther. **12**(2), 1081–1095 (2023)
- C. Jin, J.K. Udupa, L. Zhao, Y. Tong, D. Odhner, G. Pednekar, D.A. Torigian, Object recognition in medical images via anatomy-guided deep learning. Med. Image Anal. 81, 102527 (2022)
- 9. A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in Advances in Neural Information Processing Systems, vol. 25 (2012)
- 10. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp. 770–778
- G. Huang, Z. Liu, L Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2017), pp. 4700–4708
- F.C. Ghesu, B. Georgescu, Y. Zheng, S. Grbic, A. Maier, J. Hornegger, D. Comaniciu, Multi-scale deep reinforcement learning for real-time 3D-landmark detection in CT scans. IEEE Trans. Pattern Anal. Mach. Intell. 41(1), 176–189 (2017)
- H. Wen, K. Song, L. Huang, H. Wang, J. Wang, Y. Yan, Hierarchical two-stage modal fusion for triple-modality salient object detection. Measurement **218**, 113180 (2023)
- W. Zha, L. Hu, C. Duan, Y. Li, Semi-supervised learning-based satellite remote sensing object detection method for power transmission towers. Energy Rep. 9, 15–27 (2023)
- N. Le, V.S. Rathour, K. Yamazaki, K. Luu, M. Savvides, Deep reinforcement learning in computer vision: a comprehensive survey. Artifi. Intel. Rev. 55, 2733–2819 (2022)
- F. Navarro, A. Sekuboyina, D. Waldmannstetter, J.C. Peeken, S.E. Combs, B.H. Menze, Deep reinforcement learning for organ localization in CT, in *Medical Imaging with Deep Learning* (2020), pp. 544–554
- J.N. Stember, H. Shalu, Deep reinforcement learning with automated label extraction from clinical reports accurately classifies 3D MRI brain volumes. J. Digit. Imaging 35(5), 1143–1152 (2022)
- G. Maicas, G. Carneiro, A.P. Bradley, J.C. Nascimento, I. Reid, Deep reinforcement learning for active breast lesion detection from DCE-MRI, in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2017), pp. 665–673
- 19. X. Kong, B. Xin, Y. Wang, G. Hua, Collaborative deep reinforcement learning for joint object search, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2017), pp. 1695–1704
- H. Wang, Y. Chen, M. Wu, X. Zhang, Z. Huang, W. Mao, Attentional and adversarial feature mimic for efficient object detection. Vis. Comput. 39(2), 639–650 (2023)
- H. Law, J. Deng, Cornernet: detecting objects as paired keypoints, in Proceedings of the European Conference on Computer Vision, pp. 734–750

- 22. X. Zhou, J. Zhuo, P. Krahenbuhl, Bottom-up object detection by grouping extreme and center points, in *Proceedings* of the IEEE/CFV Conference on Computer Vision and Pattern Recognition, pp. 850–859
- 23. Z. Dong, G. Li, Y. Liao, F. Wang, P. Ren, C. Qian, Centripetalnet: pursuing high-quality keypoint pairs for object detection, in *Proceedings of the IEEE/CFV Conference on Computer Vision and Pattern Recognition* (2020), pp. 10519–10528
- 24. Y. Song, P. Zhang, W. Huang, Y. Zha, T. You, Y. Zhang, Object detection based on cortex hierarchical activation in border sensitive mechanism and classification-Glou joint representation. Pattern Recognit. **137**, 109278 (2023)
- K. Tong, Y. Wu, Deep learning-based detection from the perspective of small or tiny objects: a survey. Image Vis. Comput. 123, 104471 (2022)
- V. Mnih, K. Kavukcuoglu, D. Silver, A.A. Rusu, J. Veness, M.G. Bellemare, D. Hassabis, Human-level control through deep reinforcement learning. Nature 518(7540), 529–533 (2015)
- 27. C.J. Watkins, P. Dayan, Q-learning. Mach. Learn. 8, 279–292 (1992)
- 28. J. Foerster, I.A. Assael, N. De Freitas, S. Whiteson, Learning to communicate with deep multi-agent reinforcement learning, in *Advances in Neural Information Processing Systems* (2016), p. 29
- 29. J.K. Gupta, M. Egorov, M. Kochenderfer, Cooperative multi-agent control using deep reinforcement learning, in International Conference on Autonomous Agents and Multiagent Systems, pp. 66–83
- A. Vlontzos, A. Alansary, K. Kamnitsas, D. Rueckert, B. Kainz. Multiple landmark detection using multi-agent reinforcement learning, in *Medical Image Computing and Computer Assisted Intervention* (2019), pp. 262–270
- Z. Tian, L. Liu, Z. Zhang, B. Fei, Superpixel-based segmentation for 3D prostate MR images. IEEE Trans. Med. Imaging 35(3), 791–801 (2015)
- J.I. Orlando, H. Fu, J.B. Breda, K. Van Keer, D.R. Bathula, A. Diaz-Pinto, Bogunovi? H, Refuge challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs. Med. Image Anal. 59, 101570 (2020)
- X. Xu, F. Zhou, B. Liu, D. Fu, X. Bai, Efficient multiple organ localization in CT image using 3D region proposal network. IEEE Trans. Med. Imaging 38(8), 1885–1898 (2019)
- G.E. Humpire-Mamani, A.A.A., Setio B. Van Ginneken, C. Jacobs, Efficient organ localization using multi-label convolutional neural networks in thorax-abdomen CT scans. Phys. Med. Biol. 63(8), 085003 (2018)
- 35. S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: towards real-time object detection with region proposal networks, in Advances in Neural Information Processing Systems, vol. 28 (2015)
- W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.Y. Fu, A.C. Berg, Ssd: single shot multibox detector, in *European* Conference on Computer Vision (2016), pp. 21–37
- 37. J. Redmon, A. Farhadi, Yolov3: An incremental improvement (2018). arXiv preprint arXiv:1804.02767
- X. Zhu, W. Su, L. Lu, B. Li, X. Wang, J. Dai, Deformable detr: deformable transformers for end-to-end object detection (2020). arXiv preprint arXiv:2010.04159
- Z. Tian, X. Si, Y. Zheng, Z. Chen, X. Li, Multi-step medical image segmentation based on reinforcement learning. J. Ambient Intell. Hum. Comput. 13, 5011–5022 (2022)
- A. Bochkovskiy, C.Y. Wang, H.Y.M. Liao, Yolov4: Optimal speed and accuracy of object detection (2020). arXiv preprint arXiv:2004.10934

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ► Convenient online submission
- Rigorous peer review
- ► Open access: articles freely available online
- ► High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at > springeropen.com