

Virtual sports interactive system design integrating ghost net network and improved YOLOv5 algorithm

Yan Li*

Sports institute, Ningxia Normal University, Guyuan 756099, People's Republic of China

Received: 14 September 2023 / Accepted: 20 August 2024

Abstract. With the development of virtual reality, the human–computer interaction through virtual sports is gradually maturing, and users are gradually looking to interact with the two-dimensional world. The research on this type of algorithm has gained attention. However, due to the delay of the old transmission technology in the transmission of pictures, which is higher than the reaction time of human brain, the pictures are inconsistent and illogical, and the user interaction experience is poor. To solve it, this research realizes the fusion of ghost network and You Only Look Once version 5, and the simulation experiment is carried out on the data set. Firstly, the convolution block attention module is inserted into the You Only Look Once version 5 algorithm to optimize its way of calculating Hadamard product. Then, the improved algorithm and the ghost network generation fusion algorithm are combined through the direct channel. Then the algorithm is combined with the virtual sports interactive system to upgrade its key point rearrangement mode. Finally, the performance of the system is characterized on Javelin dataset, and the stability is compared with the other three algorithms. The average score of the six experiments of the system is 9.5, and the average scores of You Only Look Once version 5, ghost network and particle swarm optimization algorithm are 9.42, 9.28 and 9.36, respectively. Results show that this model has excellent performance in adjusting data volatility, and is extensive in virtual sports interaction, which can effectively improve the user experience.

Keywords: Ghost Net network / YOLOv5 algorithm / virtual reality technology / sports interactive system

1 Introduction

Recently, metaverse has risen in the wave of technological change and become a new landmark of the Internet. Virtual sports has gradually attracted more and more attention [1,2]. Virtual sports combines modern computer science and technology with traditional sports. Traditional sports are often limited by the conditions of venues and equipment. However, using the virtual sports interactive system, people can get rid of these restrictions and exercise and fitness more scientifically [3]. At present, the development of virtual sports is facing new opportunities under the new crown epidemic, and there are many ways to realize virtual sports. Although the algorithm performance based on virtual sports interactive system has been greatly improved. However, these algorithm models are complex and computationally expensive, which is not convenient for the deployment of edge devices. Convolutional neural network is used in image recognition and classification for its excellent self-learning ability, fault tolerance and

amazing operation speed [4,5]. As artificial intelligence develops, convolutional network has also been more and more applied. At present, convolutional neural networks (CNN), faster RCNN (regions with CNN features) and YOLO (You Only Look Once) series are three widely-used types. They are applied in different fields, such as video analysis, computer vision, medical image processing and so on [6]. Compared with other convolutional neural networks, the YOLO series performs well in extracting the global information of images, and also has the end-to-end training ability. In the YOLO series, the YOLOv5 algorithm stands out for its excellent convolutional neural network performance and excellent average accuracy [7]. Research on the virtual sports interaction system, combined with the convolutional neural network YOLOv5 algorithm, can be quickly and accurately deployed in edge devices [8]. Therefore, the application of YOLOv5 algorithm in virtual sports interactive system has more prominent advantages, which promotes the rapid development of virtual sports interactive system and injects new vitality and power into it. The research aims to improve system stability and optimize training ability to bring better feelings to experimenter. This research has four

* e-mail: 82007018@nxnu.edu.cn

parts. The first part is to analyze the research status of YOLOv5 algorithm and virtual reality system. The second part is to propose a virtual sports interactive system integrating the improved YOLOv5 algorithm and ghost net network. The third part is to test and analyze performance of proposed algorithm and virtual sports interactive system. The fourth part is the conclusion.

2 Related works

YOLOv5 is a deep learning algorithm for object detection. YOLOv5 is the latest member of the YOLO series. It has several novel frameworks. Many experts have made some outstanding achievements in YOLOv5 algorithm. In order to detect hidden objects in passive terahertz images automatically, accurately and in real time, Xu et al. proposed a special deep learning network and tested it on live passive terahertz images in different scenes. The results showed that the time consumption of the refined YOLOv5 detector with the best real-time performance and detection accuracy was close to 53 ms [9]. In order to make up for the limitations of photovoltaic module defect detection, Sha team proposed a deep learning method combining YOLOv5 and deep residual network. The empirical analysis of the method showed that the framework significantly improved the separation speed of photovoltaic array to 36 FPS, and fault detection accuracy of infrared image marked by segmentation region to 95% [10]. Jun et al. proposed a combined algorithm by YOLOv5 and hierarchical classification algorithm for automatic disassembly and recycling of electronic components on the circuit board. The performance of the algorithm was tested. The results showed a high recognition accuracy in the classification and positioning of components on the circuit board, and effect was improved by 38% [11]. Yi et al. proposed a model by YOLOv5 algorithm to detect insulator defects to ensure safe operation of insulators in power line inspection. The empirical analysis of this method showed that this method can obtain highly competitive results [12].

As digital economy develops, virtual reality and other technologies have penetrated into people's lives. Exploring the sports meta universe is also the Pioneer Road of industrial development. The virtual sports interactive system is in the momentum of rapid progress and will soon become a powerful track in the world. Researchers have made a lot of contributions to the virtual sports interactive system. Tsuji et al. proposed a virtual sports training system by virtual reality to solve low dynamic impedance performance problem of human upper limbs. System performance was tested. Experimental results showed that the system improved efficiency of dynamic impedance performance of human upper limbs by 5%, which had strong advantages [13]. In order to improve people's love for sports, Rejikumar et al. put forward the idea of transformation service research field, and made an empirical analysis of this idea. The experimental results showed that this research opened up a new path for the virtual sports interaction system, and raised the virtual sports interaction system to a new dimension, which was of

great significance [14]. In order to improve students' enthusiasm and sports ability, Zhou et al. proposed an interactive system by virtual sports. Experimental results showed that system application can obtain more effective training effect [15]. Zhigang proposed a method combining association rules with support vector machine for weak robustness and low accuracy in virtual sports assisted teaching system. The empirical analysis of the method showed that virtual sports assisted teaching had good training characteristics [16].

To sum up, many experts have researched YOLOv5 algorithm, which has made great contributions to its field. Many researchers have improved and broadened the application scope of the virtual sports interaction system, but the research on the integration of ghost net network and YOLOv5 algorithm in the virtual sports interaction system is very few, which has strong potential application value.

3 Virtual sports interactive system design based on improved YOLOv5 algorithm and ghost net network

This research combines the YOLOv5 algorithm with ghost net algorithm (GN). Firstly, the improved method based on YOLOv5 algorithm is introduced, and then the two algorithms are fused together. Finally, the virtual sports interaction system is constructed based on the fusion algorithm.

3.1 Improved method of adding receptive field of YOLOv5 algorithm

Among the YOLOv series algorithms, YOLOv5 has the characteristics of fewest parameters and deepest control model [17]. The study chose YOLOv5 algorithm instead of the current novel YOLOv8 or other versions, mainly because of its advantages in real-time, lightweight, easy deployment, wide applicability and stability. These advantages make YOLOv5 an ideal choice for constructing virtual sports interactive systems. When the YOLOv5 network detects the target, an additional algorithm inserted in the training sample checks the bounding box and makes specific adjustments to its width and height parameters based on the object. This process is called the generation of adaptive anchor box, as shown in Figure 1.

In Figure 1, before generating adaptive anchor frame, various parameters in it should be designed first, and then the predetermined anchor frame size can be obtained. According to the learning ability of YOLOv5 algorithm, the parameters of the object are qualified [18]. There is often a loss of confidence, which is expressed by formula (1).

$$f(\alpha) = -(1 - \alpha)^a * a \log(\alpha) - (\alpha)^a * (1 - \alpha) * \log(\alpha) \quad (1)$$

Formula (1) represents the loss function in the adaptive learning process of YOLOv5 algorithm. In formula (1), α value between 0 and 5 is called the focus loss super parameter. With the increase of the α value, the

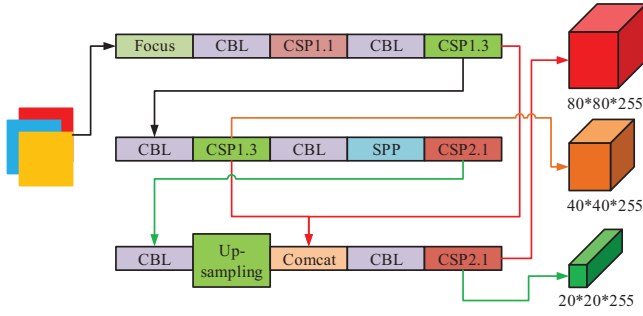


Fig. 1. Flow chart of self adaptive anchor frame generation for YOLOv5.

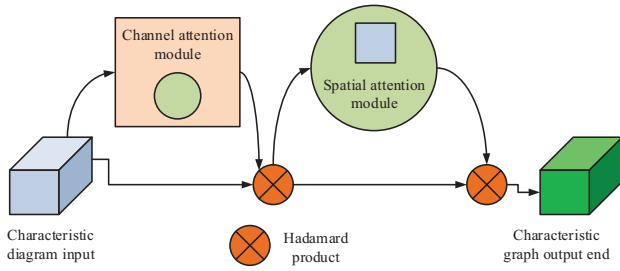


Fig. 2. Working flow chart of conflict block attention module.

significance of the loss function is gradually obvious. At this time, the model will select the loss function with higher weight for the samples with difficult classification. Adding a convolutional block attention module (CBAM) to algorithm flow can take into account channel and spatial attention mechanism in algorithm flow. CBAM process is shown in Figure 2.

The operation of Hadamard product (HP) shown in Figure 2 follows formula (2).

$$\begin{cases} F \otimes F = F' / M_C \\ F' \otimes F' = F'' / M_S \end{cases} \quad (2)$$

In formula (2), the source image is recorded as F , and the processing operations based on channel and spatial attention mechanism are recorded as M_C and M_S respectively. In the channel attention module, length, width and height of source image are W, H, C respectively set to make it pass through the feature surface with a maximum pool of dimensions. Then expected image size is set as W, H, C to make convolution input image consistent with the spatial dimension, as shown in the following formula (3).

$$M_C(F) = \zeta\{MLP[AvePool(F)] + MLP[MaxPool(F)]\}. \quad (3)$$

In the above formula (3), ζ represents channel activation function, perceptron function of multi-layer channel is recorded as MLP , and the average and maximum

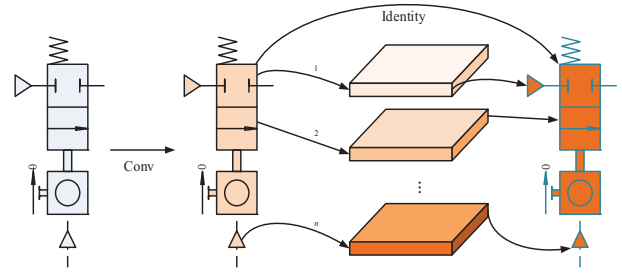


Fig. 3. Working flow chart of ghost network module.

pool in the channel are recorded as $AvePool(F), MaxPool(F)$ [19] respectively. In the spatial attention module, the image with the size of W, H, N after channel processing is used as the spatial original image, and it is made to cross the space with the size of $1, 1, C$. The image weight is adjusted through the channel to obtain the final output image, as shown in formula (4).

$$M_S(F) = \sigma^{9*9}\{[MaxP(F); AveP(F)]\}. \quad (4)$$

In formula (4), the operation of $9*9$ convolution kernel is recorded as σ^{9*9} , and the maximum and average pooling of channel modules are recorded as $MaxP(F)$ and $AveP(F)$ respectively. In order to increase the receptive field of the YOLOv5 algorithm, all positions of the weighted source image are studied and compared with the output image to obtain the long-range dependence as shown in formula (5).

$$y_i = C^{-1}(x) \sum_{A_j} \beta(x_i, x_j) g(x_j). \quad (5)$$

In formula (5), the input position of the source image is marked as i, j represents all possible positions of the image when traversing the whole world. The similarity function between the two is represented by β . x_j has the characteristics of the source image, its calculation function is marked as g . y represents the expected output characteristics, and its standardized parameter is marked as $C(x)$.

3.2 Fusion method of ghost network and improved YOLOv5 algorithm

The output image of YOLOv5 algorithm usually contains a lot of useless information. Even if the receptive field of YOLOv5 algorithm is increased, the output image still has similarities. In order to reduce this error to an acceptable range, this study proposes ghost net (GN), as shown in Figure 3 [20].

Figure 3 shows the workflow of the GN module. First, a small part of the intrinsic features of the source image is convoluted, and then the features are processed according to the complex transformation method. The two-phase fusion can mosaic the output results. The output method of a single image has a limited range of changes, but linear transformation is easy to occur when multiple images work at the same time. Therefore, this study generates targeted features based on images to output multiple results, as

shown in formula (6).

$$\phi_{ij} = \varphi_n(\phi_i'), i \in (1, n), j \in (1, s). \quad (6)$$

In formula (6), the i layer feature of a layer is denoted as ϕ_i' , and the layer with GN feature is denoted as ϕ_{ij} , φ_n can be used to retain the inner layer. The theory has some complexity and is guided by the mapping of GN modules. Set the parameters of this mapping to be the same as the network size. At this time, the compression ratio of the parameters is included in the stacking model. Since it meets the requirements of direct channel, the step size of GN module is set as 2 in this study, and the improved activation function is shown in formula (7).

$$T(x) = \begin{cases} 0 & x \in (-\infty, 0) \\ 1 & x \in (0, +\infty) \end{cases}. \quad (7)$$

It can be seen from formula (7) that when input value is negative, value of output and gradient is 0. When input value is in the positive range, the value of both is 1. It shows that the activation function easily falls into local optimum when its value is negative, which affects the iteration of neurons and prolongs the convergence time. To solve it, this study establishes a system (YOLOv5GN) that integrates YOLOv5 algorithm and GN, and takes into account the joint action of the underlying and activation functions to calculate their mean and variance, as shown in formula (8).

$$\begin{cases} \mu_p = n^{-1} \sum_{i=1}^n x_i \\ v_p^2 = n^{-1} \sum_{i=1}^n (x_i - \mu_p)^2 \end{cases}. \quad (8)$$

In formula (8), μ_p, v_p^2 are the calculations of mean and variance, p is a data set containing all x values. In the fitting process of the system, bottom layer is to learn composition of the data in training set [21]. When there is a significant difference between the two, it indicates that the system normalization ability is weak, and output channel of sample needs to be normalized, as shown in formula (9).

$$\begin{cases} \chi_i = (\chi_i - \mu_p) / (v_p^2 + \delta)^{0.5} \\ \gamma_i = \eta \chi_i + \iota \end{cases}. \quad (9)$$

In formula (9), the standardized coefficient of the sample is recorded as χ_i , η, ι is the hierarchical training parameter of the picture. When the mean value of the bottom layer of the image is 0, the smoothness of the activation function curve increases, and the decrease of monotonicity also means that the function has diversity. When the variance of this layer is 1, the sudden change of gradient will be alleviated, and the gradient existing in the negative interval will disappear. At this time, the activation function is upgraded to HSwish function, and its expression is shown in formula (10).

$$\begin{aligned} \lambda(x) &= [x \text{ReLU6}(x+3)]/6 \\ &= x[\min(\max(0, x+3), 6)]/6. \end{aligned} \quad (10)$$

In formula (10), the original activation function is written as ReLU . The HSwish function can realize the breakthrough of the operation process by improving the system hardware. The trend of the image is similar to that of the original function, but it has been improved in the process of quantitative calculation. While the action types of the two are similar, the HSwish function has a higher speed in the actual operation, and has the ability to accelerate the convergence process of the algorithm. In order to solve the local optimal problem caused by the activation function in YOLOv5 algorithm, YOLOv5GN improves the activation function and introduces the HSwish function. The new activation function improves the operation speed while maintaining the trend of the original function, and helps to speed up the convergence process of the algorithm. In addition, YOLOv5GN enhances the normalization ability of the system by calculating the mean and variance under the joint action of the bottom layer and the activation function. This helps to reduce the difference of sample output channels and improve the generalization ability and stability of the algorithm.

3.3 Virtual sports interactive system design based on fusion algorithm

Based on the YOLOv5GN fusion algorithm, in order to realize the interaction between human and machine at the virtual sports level, the study combines the YOLOv5GN model with the traditional sports interaction system (YOLOv5GN-V). Compared with the YOLOv5GN-V algorithm, it is not only superior in computing speed, generalization ability and stability, but also suitable for traditional sports interactive systems. In the YOLOv5GN-V algorithm, a feature extraction module is first added to the training system, which can simultaneously refine the confidence and affinity between key points, as shown in formula (11).

$$\begin{cases} S^I = \varpi^I(F, S^{I-1}, L^{I-1}) & \text{while } t \geq 2 \\ L^C = \varpi^C(F, S^{C-1}, L^{C-1}) & \text{while } t \geq 2 \end{cases}. \quad (11)$$

In formula (11), there are I points in S to represent the key parts of human body based on thermodynamics, and C points in L to represent the affinity of human body dynamics [22]. The two branches of the current phase are recorded as ϖ^I, ϖ^C . Steel gun projection is a common item in sports. This study can build a virtual sports interactive system of steel gun projection based on human dynamics and thermodynamics, as shown in Figure 4.

Figure 4 is schematic diagram of human-computer interaction system by steel gun projection, which is composed of camera, host, interaction plane and the person who projects the steel gun. In the human-computer interaction system, the data of human movement, interactive plane dynamic data and game video feedback data are mainly collected to ensure the effective operation of the system and the smooth experience of users.

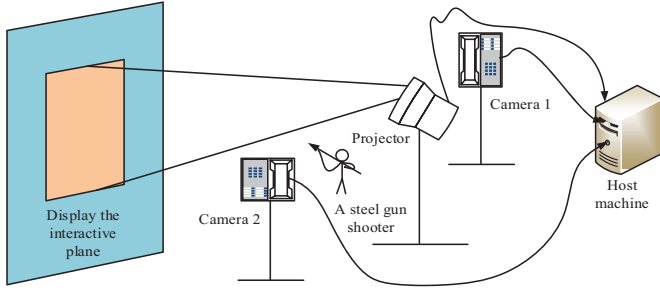


Fig. 4. Schematic diagram of man machine interaction system for steel gun projection.

The human movement data was captured by camera 2, which was used to capture and analyze the participants' movements for matching and evaluation with the YOLOv5GN-V model. Interactive plane dynamic data is recorded in real time by camera 1, which is used to monitor the activity on the interactive plane (such as the screen), assist the system in fault diagnosis and identification of user interaction status. Game video feedback data refers to the recording of the video content projected by the projector on the interactive plane and its changes, including the visual feedback when the clearance is successful, and is used to evaluate the user experience and interaction effect. The instrument configuration and data collection method is as follows: Camera 1 and camera 2 are precisely mounted and calibrated to ensure that they accurately capture the desired scene. Camera 1 is aimed at the interactive plane, while camera 2 focuses on capturing the participant's full-body movements. The projector and the host are connected by 20 M optical fiber to ensure the stable transmission of high-definition video signals to the interactive plane, while ensuring the speed and quality of data transmission. The collected video data is processed by the host, the YOLOv5GN-V model is used to analyze the human body movements, and according to the preset rules, the interaction is judged to be successful, and then the progress of the game is controlled. Based on the analysis results, the system adjusts the video content of the projector output in real time, providing the interaction with immediate visual feedback and game progress information. In this process, because the human action adopts fuzzy judgment, that is, the action amplitude can be successful within the set range, so system loss function is established in formula (12).

$$\begin{cases} \theta_S^I = \sum_{i=1}^I W(p) * \|S_i^I(p) - S_i^S(p)\|_2^2 \\ \theta_L^I = \sum_{j=1}^J W(p) * \|L_j^L(p) - L_j^C(p)\|_2^2 \end{cases} \quad (12)$$

In formula (12), the pixels of the human body image are recorded as p , at time $W(p) = 0$, the human body is not within the shooting range. When the confidence value of the background is 1, $S_i^S(p)$ indicates Branch 1. $L_j^C(p)$ represents branch 2. For the space rectangular coordinate system, the human body can be simplified as a key point, where there is a sudden change in the confidence of the

image, which is represented by formula (13).

$$S_i^I(p) = \exp\left[-\left(\|p - x_i\|_2^2\right)/\vartheta^2\right] \quad (13)$$

In formula (13), x_i is character key point position with a value range of R^2 , and the sudden change degree of confidence is ϑ . When the confidence graphs of the characters are marked, the maximum position of the pixels can be used to build the network prediction, as shown in formula (14) below.

$$S_i^S(p) = \max S_i^I(p) \quad (14)$$

In the above formula (14), the maximum value of the function is recorded as max. When character key points coordinates are all intercepted, it is necessary to rearrange them to form a simplified human body. In order to distinguish different parts of the human body and establish relationship between them, two pairs of them are set up based on the symmetry of the human body. The projection interaction process after pairing is shown in Figure 5.

Figure 5 shows flow chart of human-computer interaction system of steel gun projection game, which is composed of four processes: pretreatment, detection, human posture recognition and human-computer interaction of projection. Firstly, the image projection of the camera is split to obtain the mirror symmetrical boundary points. Then, according to the results of camera 1, the objects in the network are marked with YOLOv5GN-V model and detected and located. Then input the image of another camera, and mark the human feature points through the trained network. Finally, rearrange the feature points of the object, observe whether it conforms to the preset action, and import the conforming part into the next interaction. The correlation between candidate feature points in this process is expressed by formula (15) [23].

$$\begin{cases} E = \int_{u=0}^{u=1} L_C[p(u)] * (d_{j2} - d_{j1}) / \|d_{j2} - d_{j1}\|_2 du \\ p(u) = (1 - u)d_{j1} + ud_{j2} \end{cases} \quad (15)$$

In formula (15), two feature points d_{j1}, d_{j2} are randomly selected, and the position difference between them is recorded as $p(u)$. In the human-computer interaction system flow of the steel gun projection game shown in Figure 5, the parameters of gestures mainly include the position of feature points, the relative relationship between feature points, and the motion trajectory of feature points. The position of the feature points is used to recognize the human body posture by the YOLOv5GN-V model, and the position of each marked feature point is the key parameter. This positional information is used to describe the shape and spatial layout of the gesture. In addition to the position of individual feature points, the relative position relationship between feature points is also an important parameter. This includes the distance between feature points, angles, etc., which together define the geometric properties of the gesture. In addition, in dynamic gesture recognition, the motion trajectory of feature points over time is also an

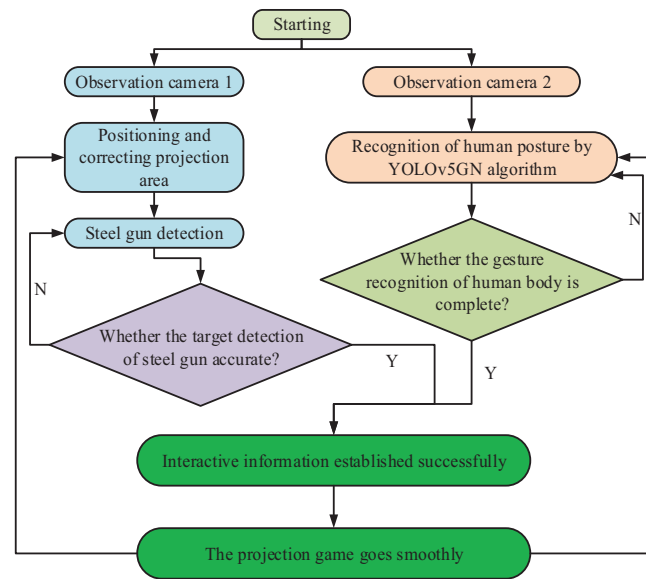


Fig. 5. Man machine interaction process of steel gun projection.

important parameter. This trajectory information reflects the changing process of gestures and is crucial for recognizing complex or continuous gestures.

4 Experimental verification of fusion algorithm combined with virtual sports interactive system

To verify analysis function of YOLOv5GN algorithm in virtual sports interaction system, this study establishes the YOLOv5GN-V model based on the fusion of the two, and verifies its iteration and accuracy. Finally, simulation experiment is conducted on Javelin data set.

4.1 Integrating ghost net network and improved YOLOv5 algorithm performance analysis

The research is verified by experiments on the javelin data set collected by ourselves, and the data set is divided into training set and test set in a ratio of 3:4. The research equipment and software used in the experiment are shown in Table 1.

To improve YOLOv5GN-V attention to key points in the picture, the attention mechanisms with different weight ratios are allocated to all levels in this study. After the GKD module is added and upgraded, the floating-point operation ability of the model is improved, which can be used to map the model. The accuracy and accuracy of this study were tested based on data set, and frame rate was 257. With gradual reduction of parameters and calculations, model complexity decreases, as shown in Figure 6.

Figure 6 shows the new era of YOLOv5GN-V model, the relationship between recall rate and accuracy rate, and compares it with YOLOv5 algorithm, gn and particle swarm optimization algorithm (PSO) algorithm to verify the superiority of the model. YOLOv5GN-V model was

Table 1. Experimental parameters.

Device type	Operating parameters or software
Operating system	Ubuntu 28.2264
Computer programming language version	3.62
Deep learning framework	Pytorch 3.14
GPU	Geforce RTX 2020ti
CPU	Intel golden 2620@3.50GHz
Language	Easy Chinese
Input size of the picture	128*128
Rate of learning	0.8
Optimizer	GKD
Computer chip	Intel I8 CPU
Size of batch	81
Data set	Javelin

slightly lower than PSO before 125 era. After 125 era, YOLOv5GN-V model always had the highest accuracy among the four algorithms in Figure 6a. In Figure 6b, as recall rate gradually increases, the accuracy performance of the four algorithms decreases. Among them, the decline rate of YOLOv5GN-V model is the slowest. It can be seen that YOLOv5GN-V has the best performance in the new era and recall rate among the four algorithms. The range parameter of pixels in the picture is characterized by receptive field. This parameter is affected by F1 value and iteration error, and their change trend is shown in Figure 7.

In Figure 7, as iteration increases, F1 values of the four algorithms show a downward trend, and the decline speed

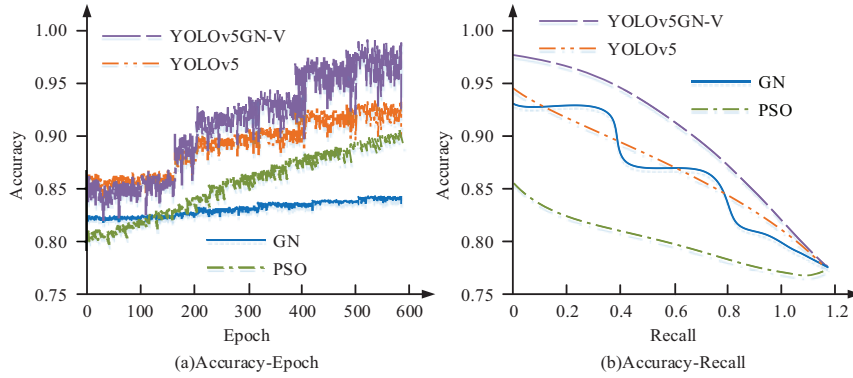


Fig. 6. Epoch and recall chart of four algorithms.

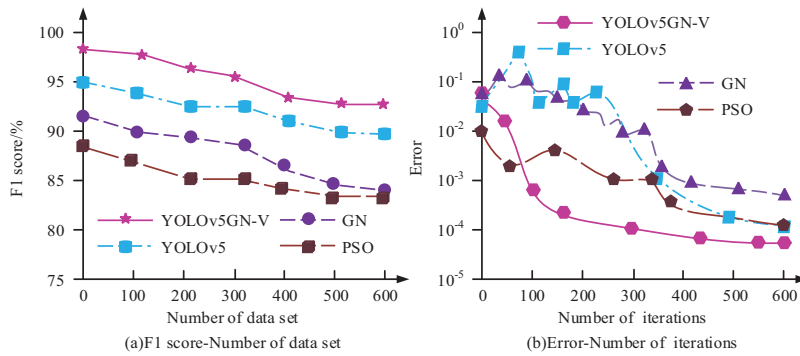


Fig. 7. Variation of F1 value and error of four algorithms with iteration times.

of YOLOv5GN-V model is the slowest. When iteration reaches 425, F1 value and iteration error of YOLOv5GN-V model show a stable trend, and their decline rate with iterations is not obvious. In Figure 7b, although PSO error before 85 iterations is lower than that of YOLOv5GN-V model, the error of PSO algorithm after that is always higher than that of the model proposed in the study, and the error of YOLOv5GN-V model in the four algorithms is the lowest, which is 7.2×10^{-5} . Considering the accuracy and cost of sports interactive experiment, the final design iteration number of this study is 425. In the iteration error, the error can be divided into three parts, including root mean square, average and maximum error. In this study, three of them were tested and analyzed, and the images were drawn as shown in Figure 8.

In Figure 8, three errors of four algorithms will also decrease as participation ratio of the training set increases. When this value reaches 52%, the average error of YOLOv5GN-V model starts to be lower than GN, and the average error after that is the lowest in Figure 8b. Figure 8c shows maximum error of the four algorithms. YOLOv5GN-V model maximum error is the best among four algorithms. To explore algorithm calculation ability in practical work, the relationship between sensitivity and specificity is established by YOLOv5GN-V, and the drawn image is shown in Figure 9.

Figure 9 shows ROC and PR curves. With specificity enhancement of the four algorithms, the sensitivity increases first fast and then slow. Algorithm strain capacity is determined by the integration between perfect predicted value and four curves. Figure 9a shows that YOLOv5GN-V model performs best among the four algorithms. In Figure 9b, the adaptability of the area judgment algorithm based on the accurate recall image and the horizontal coordinate axis is studied. Results show that YOLOv5GN-V has outstanding adaptability and can adapt to all kinds of extreme weather.

4.2 YOLOv5GN algorithm application effect in virtual sports interactive system

To verify YOLOv5GN-V effect in virtual sports interactive system, the parameters required for the experiment were set and the system was stably arranged. Among the various supplies required in the experiment, the definition of the camera and projector affects the blur degree of the picture transmitted to the host. If the picture is too vague, it will cause unclear resolution of host and affect convergence speed. To solve it, high lumen cameras and high frame rate projectors are selected in this study. Their parameters are shown in Table 2.

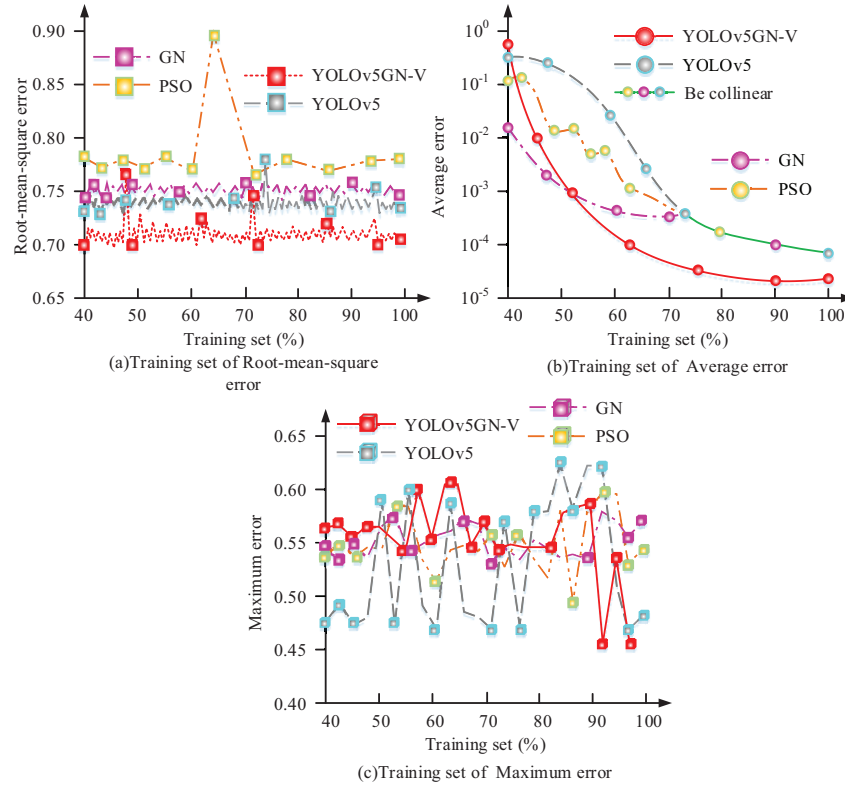


Fig. 8. Three error changes of four algorithms.

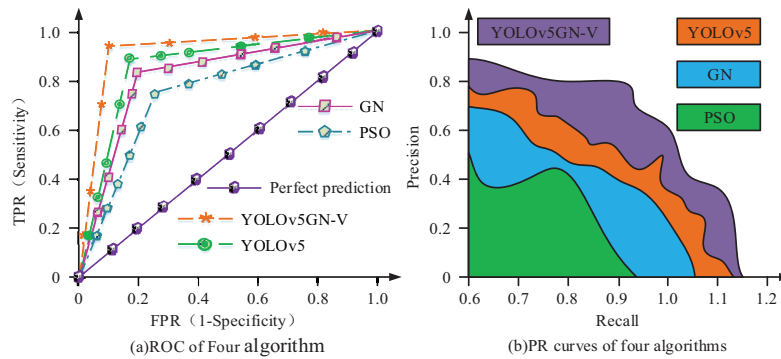


Fig. 9. ROC and PR curves of four algorithms.

Table 2. Parameters of camera and projector.

Set the category	Execution parameter
Camera photosensitive element	Doughnut
Camera pixel	More than 30 million
Maximum resolution	3000*3200
Visual angle	360 °
Imaging distance	300000 feet
Camera frame rate	100 Hz
Projector technology	Viek-d360wh
Adaptive brightness	5000 LM
Resolution ratio	512*128
Projector contrast	300000
Power efficiency	25 lm/w
Measure	12.5*11.2*13.8inches

This research is based on the virtual sports interaction system established by the YOLOv5gn model, and learns various postures of steel gun projection, which is used to evaluate the human posture in the experiment, and then score its accuracy. The main interface of the system consists of a steel gun recognition algorithm and a human feature point delineation algorithm, which can judge whether it meets the standard through the scores of experimenters in different environments, such as flat land, mountains, oceans, sand dunes and so on. To increase shooter accuracy, an adaptive wild monster with terrain generation is added to the system. The experimenter is scored by distance between javelin and target centroid. To evaluate the universality of the algorithm, the research conducted several experiments on the shooting scores of the four algorithms, and the drawing is shown in Figure 10.

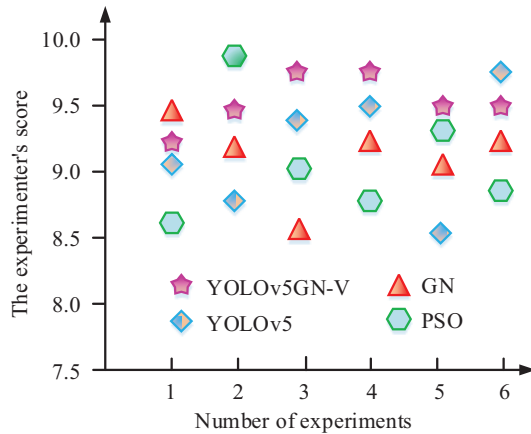


Fig. 10. Experimental results of extensive testing of the algorithm.

In the six experiments in Figure 10, the shooting scores of the YOLOv5GN model were 9.22, 9.45, 9.72, 9.72, 9.48 and 9.48 respectively. The small fluctuation range proved that there were no abnormal points. Even if the initial score was low, it would steadily improve under the correction of the YOLOv5GN-V system. The average score of six experiments is 9.50, which shows that the model has superior performance in adjusting data volatility. The mean scores of YOLOv5, GN and PSO algorithms were 9.42, 9.28 and 9.36, respectively, and the fluctuation range was large. Compared with them, the stability and training ability of the YOLOv5GN-V system are the best, which shows that the model is suitable for human-computer interaction of virtual sports and can bring better feelings for the participants.

5 Conclusion

As human-computer interaction of virtual sports system progresses, users' requirements for the authenticity of playing are gradually improved, and the interactive algorithm is upgraded. This study uses convolution block attention module to optimize the YOLOv5 algorithm and integrate it with GN. It is found that the fusion algorithm can not only expand the definition domain, but also quickly jump out of the local optimum. The Hadamard product and direct channel of algorithm are studied and simulated on Javelin data set. Through the test experiment, it is found that the final era value is 125, and the model error is the highest when iteration is 425. F1 value, root mean square error and average maximum error of the model are also judged and plotted based on this

PR curve and ROC curve. Results show that YOLOv5GN-V performance is the best. To verify model stability and universality, the score verification on the steel gun projection movement was studied, and compared with YOLOv5, GN and PSO. For the YOLOv5GN-V system, the scores of six experiments are 9.22, 9.45, 9.72, 9.72, 9.48 and 9.48 respectively, and the average value is recorded as 9.50, while the experimental results of the other three algorithms are 9.42, 9.28 and 9.36 respectively. In addition, in the six experiments, the data fluctuation range of YOLOv5GN-V

system is small, while the fluctuation range of YOLOv5, GN and PSO is large. It shows that YOLOv5GN-V system has high rigidity and can adapt to various extreme situations. However, this study only detects the projective steel gun, and there are still many sports. And the path conditions are diverse, and the research is only aimed at the desert and other four kinds of terrain. It indicates that the research scope of this study is relatively narrow, because data are limited. With the increase of data, future research will make breakthrough progress.

Funding

This research received no external funding.

Conflicts of interest

The author reports there are no competing interests to declare.

Data availability statement

Data will be made available on reasonable request.

Author contribution statement

All work for this article was completed by Yan Li.

References

1. P.K. Yadav, J.A. Thompson, S.W. Searcy, Assessing the performance of YOLOv5 algorithm for detecting volunteer cotton plants in corn fields at three different growth stages, *Agric. Artistic Intell.* **6**, 292–303 (2022)
2. Y. Zhang, K. Guo, Power plant indicator light detection system based on improved YOLOv5, *J. Beijing Inst. Technol.* **31**, 605–612 (2022)
3. J. Zhang, Y.B. Hu, J.L. Yang, Automatic counting of retinal ganglion cells in the entire mouse retina based on improved YOLOv5, *Zoolog. Res.* **43**, 738–749 (2022)
4. A. Singh, M.M. Gupta, Convergence of machine learning and statistics to predict COVID-19 evolution, *Acta Inform. Malaysia* **6**, 34–38 (2022)
5. A.M. Thantawi, S.A. Indriyati, Conceptual design impacts in new normal era: the use of artificial intelligence (AI) And Internet of Things (IOT) (case studies: class room and restaurant), *Acta Inform. Malaysia* **6**, 39–42 (2022)
6. K. Karunamurthy, M.M. Feroskhan, G. Suganya, I. Saleel, Prediction and optimization of performance and emission characteristics of a dual fuel engine using machine learning, *Int. J. Simul. Multidisci. Des. Optim.* **13**, 13 (2022)
7. D. Bassir, H. Lodge, H. Chang, J. Majak, G. Chen, Application of artificial intelligence and machine learning for BIM: review, *Int. J. Simul. Multidisci. Des. Optim.* **14**, 5 (2023)
8. L. Xu, S. Dong, H. Wei, Q. Ren, J. Huang, J. Liu, Defect signal intelligent recognition of weld radiographs based on YOLO V5-improvement, *J. Manufactur. Processes* **99**, 373–381 (2023)

9. F. Xu, X. Huang, Q. Wu, X. Zhang, Z. Shang, Y. Zhang, YOLO Msf: toward real time detection of compromised objects in passive terahertz images, *IEEE Sens. J.* **22**, 520–523 (2022)
10. Z. Sha, H. Feng, X. Rui, Z. Zeng, Pig tracking utilizing fiber optical distributed vibration sensor and YOLO, *J. Lightwave Technol.* **236**, 406–416 (2021)
11. C. Jun, B. Erdemt, P. Jingyu, classification and positioning of circuit board components based on improved YOLOv5, *Proc. Comput. Sci.* **38**, 613–626 (2022)
12. L. Yi, M. Ni, Y. Lu, Insulator defect detection for power grid based on light correction enhancement and YOLOv5 model, *Energy Rep.* **8**, 807–814 (2022)
13. T. Tsuji, Y. Sumida, M. Kaneko, S. Awamura, A virtual sports system for skill training, *J. Robot. Mech.* **13**, 168–175 (2021)
14. G. Rejikumar, A. Jose, S. Matthew, D.P. Chacko, A. Asokan Ajitha, Gifts a theory of well being in digital sports viewing behavior, *J. Serv. Market.* **36**, 245–263 (2022)
15. J. Zhou, Virtual reality sports auxiliary training system based on embedded system and computer technology, *Microprocess. Microsyst.* **82**, 1–6 (2021)
16. T. Zhigang, Research on decision support system of sports assistant teaching and training based on association rules and support vector machine, *J. Intell. Fuzzy Syst.* **15**, 1–12 (2021)
17. S. Oslund, C. Washington, A. So, T. Chen, H. Ji, Multi view robot advantageous stickers for Arabic objects in the physical world, *J. Comput. Cogn. Eng.* **1**, 152–158 (2022)
18. H. Li, Virtual interaction algorithm of cultural heritage based on multi feature fusion, *J. Comput. Methods Sci. Eng.* **22**, 333–347 (2022)
19. T. Liu, M. Hu, S. Ma, Exploring the effectiveness of karst interaction in driver assistance systems via virtual reality, *IEEE/CAA J. Automat. Sin.* **9**, 1520–1523 (2022)
20. H. Wang, J. Wu, A virtual reality based surgical skills training simulator for catheter ablation with real-time and robust interaction, *Virt. Real. Intell. Hardw.* **3**, 302–314 (2021)
21. W. Mai, L. Fang, Z. Chen, Application of the somatosense interaction technology combined with virtual reality technology on upper limbs function in cerebrovascular disease patients, *J. Biomed. Sci. Eng.* **13**, 66–73 (2020)
22. H. Bai, L. Zhang, J. Yang, Bringing full featured mobile phone interaction into virtual reality, *Comput. Graph.* **97**, 42–53 (2021)
23. M. Naver, M. Olgun, E. Türkarlan, Cosine and tangent similarity measures based on Choquet integral for spherical fuzzy sets and applications to pattern recognition, *J. Comput. Cogn. Eng.* **1**, 21–31 (2022)

Cite this article as: Yan Li, Virtual sports interactive system design integrating ghost net network and improved YOLOv5 algorithm, *Int. J. Simul. Multidisci. Des. Optim.* **15**, 19 (2024)