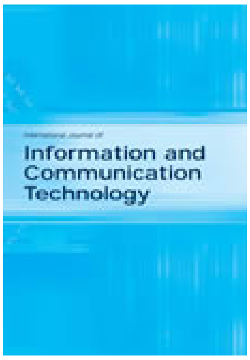


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Gomoku game recognition and localisation using image processing and deep learning

Huanxi Chen, Wei Li, Yong Shen, Xiang Ding, Shunbin Li, Minghao Gu, Dexin Sun, Zhenhua Tan, Guochao Wang

Article History:

Received:	29 September 2024
Last revised:	06 December 2024
Accepted:	06 December 2024
Published online:	28 April 2025

Gomoku game recognition and localisation using image processing and deep learning

Huanxi Chen and Wei Li*
School of Machinery and Transportation,
Southwest Forestry University,
Kunming, 650224, China
Email: 2467223709@qq.com
Email: 772074913@qq.com
*Corresponding author

Yong Shen
School of Electrical Information Engineering,
Yunnan Minzu University,
Kunming, 650504, China
Email: kmsmg@163.com

Xiang Ding, Shunbin Li, Minghao Gu,
Dexin Sun, Zhenhua Tan and Guochao Wang
School of Machinery and Transportation,
Southwest Forestry University,
Kunming, 650224, China
Email: 2434845188@qq.com
Email: 3362056303@qq.com
Email: 290789628@qq.com
Email: 1830179685@qq.com
Email: 949813510@qq.com
Email: 1453728499@qq.com

Abstract: Aiming at the accuracy of Gomoku game recognition during human-machine game, this study introduces a method for Gomoku game recognition and localisation using image processing combined with deep learning techniques, providing accurate position information for subsequent game decision-making and control driving in human-machine game devices. Firstly, the image processing technology is used to extract and correct the Gomoku board area, and then the corner points of the board are obtained by corner division. Secondly, to enhance the accuracy of Gomoku piece recognition, spatial attention (SA) is introduced into the C2f module of the backbone network of the yolov8 model to achieve the piece recognition. Finally, the location of the Gomoku piece is completed by comparing the position of the piece and the corner point. After experimental validation, the improved model shows a marked improvement over the original yolov8 model, with the precision rate, recall rate, mAP50 and mAP50-95 being 99.9%, 99.0%, 99.4% and 96.4%, respectively. Even under challenging conditions such as uneven brightness, low brightness and tilting of the camera, the recognition and position judgment of the Gomoku pieces can still maintain high accuracy. The introduced method has high performance in Gomoku piece recognition and localisation, providing favourable support for visual detection of Gomoku gaming devices.

Keywords: human-machine game; image processing; yolov8; attention mechanism; piece recognition; localisation.

Reference to this paper should be made as follows: Chen, H., Li, W., Shen, Y., Ding, X., Li, S., Gu, M., Sun, D., Tan, Z. and Wang, G. (2025) 'Gomoku game recognition and localisation using image processing and deep learning', *Int. J. Information and Communication Technology*, Vol. 26, No. 9, pp.43–63.

Biographical notes: Huanxi Chen received his Bachelor's degree in 2022 from Zhaoqing College, majoring in Mechanical Design, Manufacturing and Automation. Since September 2022, he is pursuing his Master's degree at Southwest Forestry University with a major in Mechanical Manufacturing and Automation. Currently, his main research interests include machine vision and intelligent control.

Wei Li is a Professor and the Dean of the School of Machinery and Transportation at Southwest Forestry University. She graduated from the School of Mechanical and Electrical Engineering at Kunming University of Science and Technology in 2011 with a PhD in Mechanical Design and Theory. She hosted multiple projects including Yunnan Provincial Fund Project, Yunnan Province Eight Major Brand Professional Construction Project, Yunnan Province Virtual Simulation Experiment Project, Graduate Excellent Course Project, Graduate Classic Case Library, participated in multiple National Natural Science Foundation projects, and led multiple horizontal projects. Her main research directions include mechatronics technology, intelligent manufacturing technology, industrial design, etc.

Yong Shen is a Professor and a Master's Supervisor. He graduated from Hefei University of Technology in 1988 with a Bachelor's degree. In July 2001, he joins the School of Electrical Information Engineering, Yunnan Minzu University, and the research field is industrial automation. He has won multiple awards, including the third prise for scientific and technological progress in Kunming, the second prise for scientific and technological progress in Yunnan Province, the special prise for science and technology from China Railway Construction Corporation, and the first prise for scientific and technological progress in Yunnan Province.

Xiang Ding is studying at Southwest Forestry University since September 2022, majoring in Machinery. He is a third-year engineering Master's student in the School of Mechanical Engineering and Transportation. His main research focus is on mechatronics and machine vision.

Shunbin Li graduated from Southwest Petroleum University, since September 2022, and is studying in Southwest Forestry University, majoring in Machinery. He is currently a third-year Master's degree student in the College of Mechanics and Transportation. His main research direction is mechatronics.

Minghao Gu is a graduate student of Machinery, School of Machinery and Transportation, Southwest Forestry University. His research direction includes modern manufacturing and control technology, good at electrical control and intelligent control algorithm research. He has practical foundation for CNC related work, and rich extracurricular practical experience. He previously worked at the South China Laboratory of the China Quality Certification Center. During university, he participated in class activities and served as the class organisation committee member.

Dexin Sun is majoring in Automation at Shandong Agricultural University from September 2018 to June 2022. Since September 2022, he has been studying in Southwest Forestry University, majoring in Machinery. Currently, he is a third-year professional Master student in the School of Machinery and Transportation. His main research directions are machine vision and mechatronics.

Zhenhua Tan is studying in Southwest Forestry University since September 2022, majoring in Agricultural Engineering and Information Technology. He is currently a third-year professional Master student in the School of Machinery and Transportation. His main research direction is mechatronics.

Guochao Wang is studying in Southwest Forestry University since September 2022, majoring in Mechanical Manufacturing and Automation. He is currently a third-year Master's degree student in the College of Mechanics and Transportation. His main research directions are finite element analysis and mechanical dynamics.

1 Introduction

As artificial intelligence rapidly evolves, machine gaming (Chen et al., 2020) as a research topic within the domain of artificial intelligence, it has garnered significant attention from scholars, based on reinforcement learning technology AlphaGo and other chess game artificial intelligence (Silver et al., 2016, 2017), the research of human-machine game has attracted extensive attention from the academic community. The human-machine game involves competition between humans and machines. Scholars have increasingly applied these concepts to practical gaming equipment, the development of intelligent human-machine game equipment such as chess robot (Zhang, 2022).

The rise of deep learning has produced a series of network models such as YOLO, VGG, etc. Some scholars have adopted deep learning in the fields of medicine (Prusty et al., 2024a; Mishra et al., 2023), industry (Zhu et al., 2024), mining (Yuan and Li, 2022), etc., using these models for detection and classification tasks. For instance, in industry, deep learning models are used for defect detection (Zhu et al., 2024), while in the medical field, VGG and GAN are employed for cancer (Prusty et al., 2022, 2024b) and disease (Patil et al., 2023) classification. Undoubtedly, deep learning has become a boon for various fields such as industry and medicine. The YOLO detection algorithm (Zhu et al., 2024), as a single-stage detection algorithm, is known renowned in target detection for its fast and accurate performance. Recently, the attention mechanism has been incorporated into deep learning and widely used it for target detection (Li et al., 2024) by researchers. Adding the attention mechanism enhances the model's emphasis on critical areas, which has a significant effect on the detection in a complex environment.

The working principle of the human-machine game system is as follows: first, the type and position of the pieces are obtained through the piece recognition system. Then, based on this positional information and the piece moving strategy, the position for the next move is determined. And then the position is transmitted to the control system via data conversion, which drives the actuators to complete the move. It can be seen that the recognition and localisation of pieces play a crucial role in the human-machine game system. To address the problem of chess piece localisation, some scholars have proposed a sensor-based assisted localisation method (Larregay et al., 2018), although this method can improve the accuracy, it requires a specially customised chess board. With the progress of computer vision technology, more and more intelligent devices begin to adopt visual recognition to realise their intelligent functions. In human-machine gaming systems, such as checkers gaming system (Fabris et al., 2024) and Go gaming system (Zhao, 2022), computer vision technology is widely used to automatically identify the positions of the pieces, which significantly improves the intelligence of the system.

Vision-based recognition in Gomoku is a crucial component for human-machine interaction systems. Regarding the issue of Gomoku board recognition, various scholars have proposed a series of methods. Wu et al. (2019) used edge detection and Hough transform to detect the board and applied template matching to recognise and locate the pieces. Mao et al. (2017) extracted the board's edges through edge detection, segmented the board, and then used template matching to identify the pieces and complete their localisation. Ou (2020) also used edge detection and other algorithms to extract the board, followed by combining the Hough transform and greyscale values to recognise the pieces and achieve their localisation. Zhang and Gu (2022) employed the Hough transform method with angle-radius representation to detect board lines, converting it into a slope-intercept form to detect corners and calculate intersection points to determine board corner points. They then combined greyscale values to identify the pieces and complete the localisation. Wang et al. (2021) binarised the board image to obtain board lines, reconstructed the board by calculating the height and width of rows and columns, and identified piece types based on pixel values, completing the final localisation of the pieces. Xue et al. (2023) used an edge detection method to obtain the board's contours, applied perspective transformation, and then completed the subsequent piece localisation. From the above methods, it can be summarised that board processing mainly relies on edge detection to obtain the crossing lines and contours of the board, while piece recognition is generally based on greyscale values, Hough transform, and template matching. However, in the cases of local shadows, uneven lighting, or severe image noise, edge detection may result in false contours, making it difficult to extract complete board features. For piece recognition, greyscale-based methods rely on parameter settings, and under uneven lighting or complex light sources, the instability of greyscale parameters may lead to inaccurate results. While template matching performs well under standard conditions, its accuracy decreases under varying lighting conditions, and its high computational complexity hinders real-time processing. Therefore, in complex environments with lighting variations or camera tilt, the aforementioned methods may face limitations in recognition accuracy compared to ideal constant lighting conditions. To address the challenge of uneven lighting, scholars have proposed using threshold and morphological processing techniques (Wang et al., 2020; Li and Wang, 2020) for image segmentation, which have shown promising results.

In conclusion, scholars have proposed various algorithms for board recognition in human-machine interaction systems. However, most of these methods are based on conditions such as uniform lighting and well-placed boards, with less focus on scenarios involving uneven lighting, low-light conditions, or camera tilts. Therefore, this paper introduces a Gomoku recognition and localisation method based on image processing and deep learning, specifically targeting these complex conditions. Image processing techniques are used to extract the board, followed by perspective transformation and corner point division to obtain the board's corner points. To address issues such as uneven lighting, an improved yolov8 model is proposed, integrating an attention mechanism to improve the model's capacity for recognising pieces in challenging environments. Finally, by calculating, the position of the piece is determined whether it is located at the corner of the board, thus obtaining the corresponding piece coordinates, achieving precise locating of the piece, providing real-time position information to the upper computer for each placement, and obtaining the next placement position using the piece moving strategy based on the positional data obtained by the game recognition algorithm. Subsequently, the position data is transmitted to the control system through the communication protocol to achieve accurate piece placement operation. Through this method research, the paper aims to provide insights for the development of vision systems in game-playing robots or similar control demand devices.

2 Related work

The workflow of human-machine gaming device is as follows: firstly, obtaining the information of Gomoku pieces through visual recognition system, then obtaining the position of the next move based on the information of pieces under the calculation of strategy system, and then feeding back the position of the next move to the control system to drive the actuator to complete the operation of piece placement. The whole framework is shown in Figure 1. This paper mainly carry out the research on the recognition and localisation method of Gomoku piece based on image processing and deep learning in human-machine gaming equipment.

The Gomoku board consists of a 15 × 15 grid, with a total of 225 intersection points, and the spacing between each point is uniform. There are two types of pieces in Gomoku: black and white, and the localisation of each piece depends on which intersection point it lands on. In this paper, the following three parts of the research are centred on the problem of recognising and localising the pieces during the game of Gomoku.

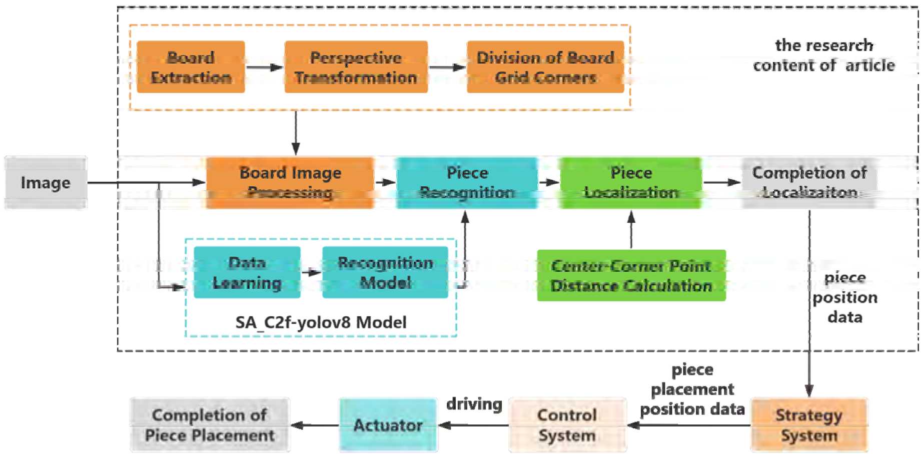
1. Perform image processing on the board to extract and divide the board grid's intersection points. Image processing techniques are used to extract the board grid area and identify the four vertices of the board grid. Then, the board is corrected to a standard shape using perspective transformation. By applying interpolation, the all intersection points of the board grid are extracted, and inverse perspective transformation is used to derive the intersection points from the original board image, thus determining the precise coordinates of the points.

2. Use deep learning to recognise the pieces. To enhance piece recognition accuracy in different lighting conditions, a dataset created from board photos taken in different lighting environments is used for deep learning, while the SA attention mechanism is integrated into the C2f module of the yolov8 model to improve feature extraction abilities. The improved yolov8 model is then used to learn from the board photos, increasing piece recognition accuracy in varying lighting conditions.

3. Piece localisation. Based on the first two steps, piece localisation is performed by determining which intersection point the piece is located on. A centre-corner point

distance calculation method is introduced: when the distance between the centre of the piece and the corner point is within the radius of the piece, the piece is considered to be located on that point. This completes the piece localisation in a Gomoku game.

Figure 1 Research framework and process (see online version for colours)



3 Board image processing

3.1 Board extraction

Board extraction is critical for accurate piece localisation. The first step involves extracting the board grid area, followed by identifying the four corners of the board grid area to facilitate subsequent board correction and division of intersection points.

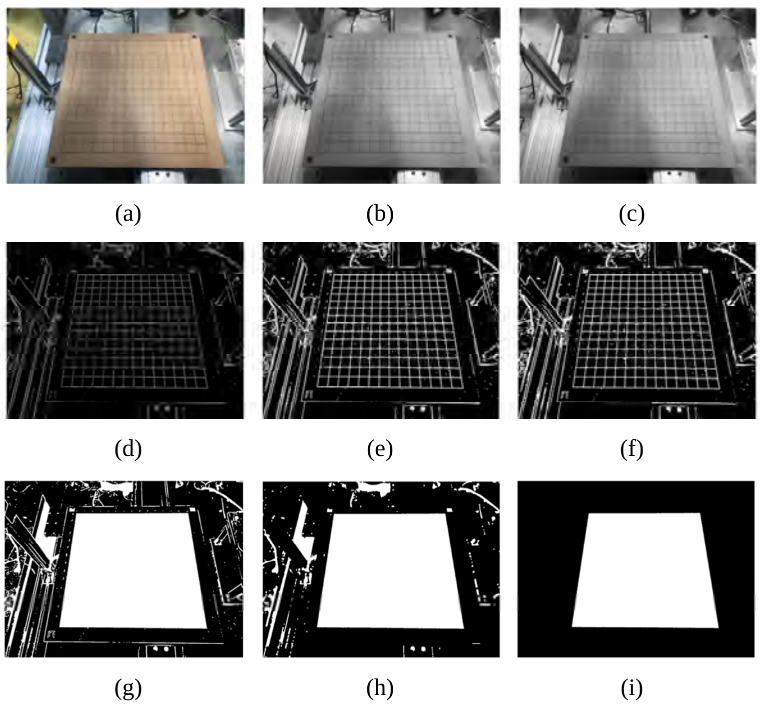
Currently, there are various methods for board extraction. Common approaches include detecting board lines using Hough line transform to extract the board area, segmenting the board area based on colour information, and applying global threshold combined with morphological processing to extract the board. Finally, the four corner points are detected using Hough line detection. To avoid the influence of the cluttered environment around the board on the positioning of the board, in the research of this paper, binarisation and morphological processing are used to directly extract the grid line region on the board, instead of extracting the outline of the whole board, so as to effectively avoid the influence of the cluttered environment around the board on the extraction of the grid line region. In addition, in the process of image processing, the outermost edge of the grid line is extracted directly to obtain the four corner points of the grid line area, and the inner corner points are obtained by interpolation, thus avoiding the situation of imprecise extraction of the corner points caused by the unclear local lines inside the grid line.

The overall image processing workflow for the board is as follows: the captured image is first subjected to greyscale conversion, Gaussian blur, adaptive binarisation, morphological processing, and connected domain analysis to obtain a complete white grid area on the board. The contour of the grid area is then directly extracted, and the four corners are identified. Figure 2(a) shows the input image, and we apply a weighted method to convert the RGB image to greyscale. The calculation formula is as follows:

$$Gray = 0.299 \times R + 0.587 \times G + 0.114 \times B \quad (1)$$

where R, G and B denote red, green, and blue, respectively.

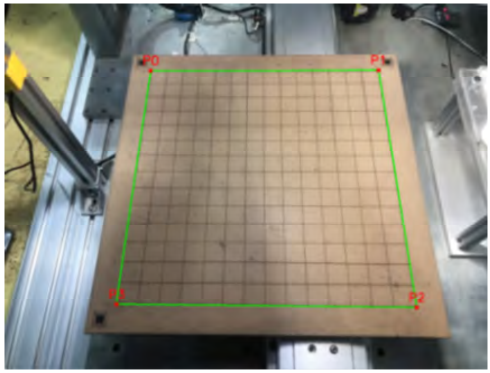
Figure 2 Board processing, (a) original image (b) greyscale the original image (c) noise removal with Gaussian blurring (d) conversion to binary map (e) inflation of the binary map; (f) closure operation to connect regions (g) hole filling (h) remove the small smudges by the open operation (i) extract the maximum connectivity domain (see online version for colours)



Greyscale image is shown in Figure 2(b). Gaussian blur is used on the greyscale image to eliminate noise, as shown in Figure 2(c). Followed by adaptive binarisation to convert it into black and white, as shown in Figure 2(d). At this stage, there are some unwanted impurities around the board area in the binary image, so morphological processing is applied. First, dilation is performed on the binary image, which makes the board's contours more distinct, as shown in Figure 2(e). Then, joining disconnected regions in binary images using the closure operation, as shown in Figure 2(f). Afterward, hole filling is used on the binary image after the closing operation, filling the board grid region with white, as shown in Figure 2(g). However, some small white specks remain in the image. To remove these small specks, an opening operation is applied, as shown in Figure 2(h). Finally, the largest connected region is extracted, leaving only the largest white area, which corresponds to the desired board grid area, as shown in Figure 2(i).

After the aforementioned image processing steps, the resulting white area corresponds to the board grid region. By extracting the contour of this white area, the four corner points of the board grid are obtained, as shown in Figure 3. The green box indicates the board area, and the four points labelled 'P' represent the four corner points of the board grid region.

Figure 3 Board region and corner detection (see online version for colours)



3.2 Board correction

Board correction usually requires identifying the four corner points of the board and performing a perspective transformation based on these points. Traditional methods for extracting the corner points of the board include using corner detectors and feature-based approaches, but these methods impose high demands on the surrounding environment and geometric features. In this study, the contour of the board grid is obtained through image processing, and the four corner points are directly extracted to complete the board correction.

Let the four corner points of the board grid be defined as $P_0(x, y)$, $P_1(x_1, y_1)$, $P_2(x_2, y_2)$ and $P_3(x_3, y_3)$, with their corresponding points after transformation being $P_0'(u, v)$, $P_1'(u_1, v_1)$, $P_2'(u_2, v_2)$ and $P_3'(u_3, v_3)$. According to the principles of perspective transformation, the relationship holds as follows.

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = H \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{2}$$

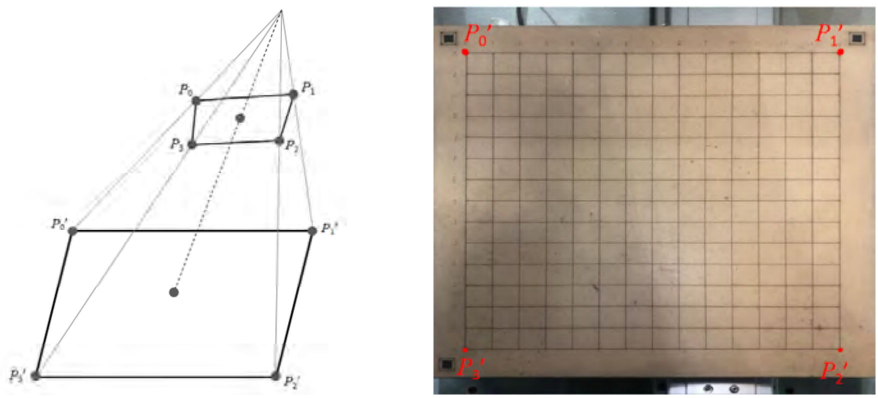
Convert the obtained coordinates of the board grid vertices into three-dimensional homogeneous coordinates under the action of a transformation matrix, and then use perspective division to divide x' and y' by z' to obtain the corresponding two-dimensional projection coordinates. Thus, from equation (2), the following can be derived.

$$\begin{cases} u = \dfrac{x'}{z'} = \dfrac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}} \\ v = \dfrac{y'}{z'} = \dfrac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}} \end{cases} \tag{3}$$

where z' is the normalisation factor, which transforms the three-dimensional homogeneous coordinates (x', y', z') used for the intermediate transformation to the target coordinates (u, v), H is the transformation matrix.

The most important thing in perspective transformation is to obtain the transformation matrix H. Since (x, y) is obtained by image processing, and (u, v) can be specified, then the H can be obtained according to equation (3) in the above case. And then the H is used in the perspective transformation to complete the projection of the board. As shown in Figure 4 for the whole process of perspective transformation, where Figure 4(b) shows the diagram of the board after perspective transformation.

Figure 4 Perspective transformation process, (a) perspective transformation (b) corrected image (see online version for colours)



(a) (b)

3.3 *Division of board grid corners*

In this study, a linear interpolation method is used to divide the 15 × 15 inter-section points of the board grid based on the four extracted corner points. The division is carried out on the perspective-transformed board image [Figure 4(b)]. The overall approach involves first dividing the line segments along the board's left and right sides, followed by further division of the points between the corresponding rows on each side.

The specific division method is as follows: first, interpolation is applied to the left and right boundaries of the board grid to calculate the intersection points $m_{left}(i)$ and $m_{right}(i)$ on the two line segments $P_0'P_3'$ and $P_1'P_2'$, respectively. Then, based on the points $m_{left}(i)$ and $m_{right}(i)$, the horizontal intersection points $m_{horizontal}(j)$ for the i^{th} row are calculated. The detailed calculation is as follows:

The vertical interpolation calculation is given as follows:

$$\begin{aligned} m_{left}(i) &= \left(1 - \frac{i}{14}\right) \times P_0' + \frac{i}{14} \times P_3' \\ m_{right}(i) &= \left(1 - \frac{i}{14}\right) \times P_1' + \frac{i}{14} \times P_2' \end{aligned} \tag{4}$$

The horizontal interpolation calculation is given as follows:

$$m_{horizontal}(j) = \left(1 - \frac{j}{14}\right) \times m_{left}(i) + \frac{j}{14} \times m_{right}(i) \tag{5}$$

where i represents the rows (0 to 14) and j represents the columns (0 to 14).

Since the above calculation is to divide the corner points of the perspective transformed chessboard grid, in order to be able to divide the corners on the original image of the board grid, it is necessary to invert the corner points of the perspective transformed board grid so as to get the coordinates of the corner points on the original board, then the transformation matrix H obtained from equation (3) can be inverted according to equation (6) to derive the inverse matrix H^{-1} of the transformation matrix.

$$H \times H^{-1} = I \tag{6}$$

According to the matrix H^{-1} the corner point $m(u, v)$ obtained from the previous calculation can be mapped to the original image by inverse transformation as follows:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = H^{-1} \times \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{7}$$

Assume that the inverse matrix H^{-1} is obtained as follows:

$$H^{-1} = \begin{bmatrix} h'_{11} & h'_{12} & h'_{13} \\ h'_{21} & h'_{22} & h'_{23} \\ h'_{31} & h'_{32} & h'_{33} \end{bmatrix} \tag{8}$$

From equations (7) and (8), the two-dimensional corner coordinates in the projection map are transformed to the three-dimensional homogeneous coordinates (x', y', z') of the corner points of the original image under the action of the inverse matrix, and the specific transformation process is shown in equation (9), so as to facilitate the subsequent transformation of the two-dimensional corner point coordinates in the original image.

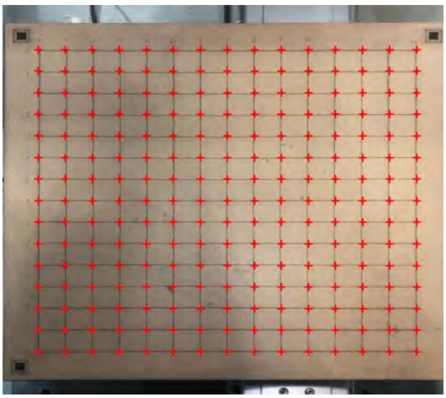
$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} h'_{11} & h'_{12} & h'_{13} \\ h'_{21} & h'_{22} & h'_{23} \\ h'_{31} & h'_{32} & h'_{33} \end{bmatrix} \times \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} h'_{11}u + h'_{12}v + h'_{13} \\ h'_{21}u + h'_{22}v + h'_{23} \\ h'_{31}u + h'_{32}v + h'_{33} \end{bmatrix} \tag{9}$$

From equation (9), the three-dimensional coordinates (x', y', z') of the corner points in the original image of the board have been obtained by inverse transformation. The coordinates of the corner points (x', y', z') are converted into coordinates (x, y) by dividing z' respectively, the specific transformation process is shown in equation (10). The (x, y) are the coordinates of the corner points in the original diagram of the board, according to these coordinates to extract the corner points of the board, to complete the preparation of the judgment of the position of the pieces.

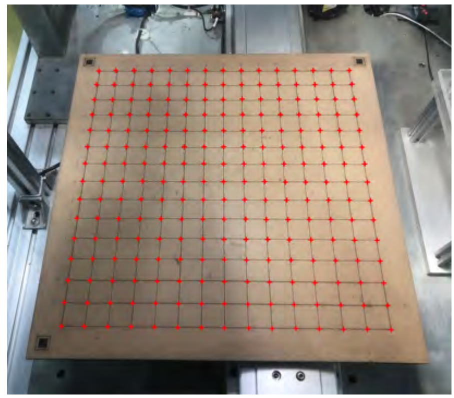
$$\begin{cases} x = \dfrac{x'}{z'} = \dfrac{h'_{11}u + h'_{12}v + h'_{13}}{h'_{31}u + h'_{32}v + h'_{33}} \\[2ex] y = \dfrac{y'}{z'} = \dfrac{h'_{21}u + h'_{22}v + h'_{23}}{h'_{31}u + h'_{32}v + h'_{33}} \end{cases} \tag{10}$$

Based on the above calculations, the corner points of the perspective-transformed board image are mapped back to the original board through an inverse transformation, thereby obtaining the corner points of the uncorrected original board, as illustrated in Figure 5. Each position of the corner point is documented for the subsequent localisation of pieces.

Figure 5 Board corner point division, (a) corrected board (b) original board (see online version for colours)



(a)



(b)

4 Piece recognition and localisation

For the problem of piece recognition, this study adopts a deep learning approach, using an improved yolov8 model to learn from the game dataset, thereby improving the accuracy of piece recognition. For the problem of piece localisation, the recognised pieces are compared with the corner points of the board. By calculating the distance between the piece and the corner points, the coordinates of the piece on the board are determined.

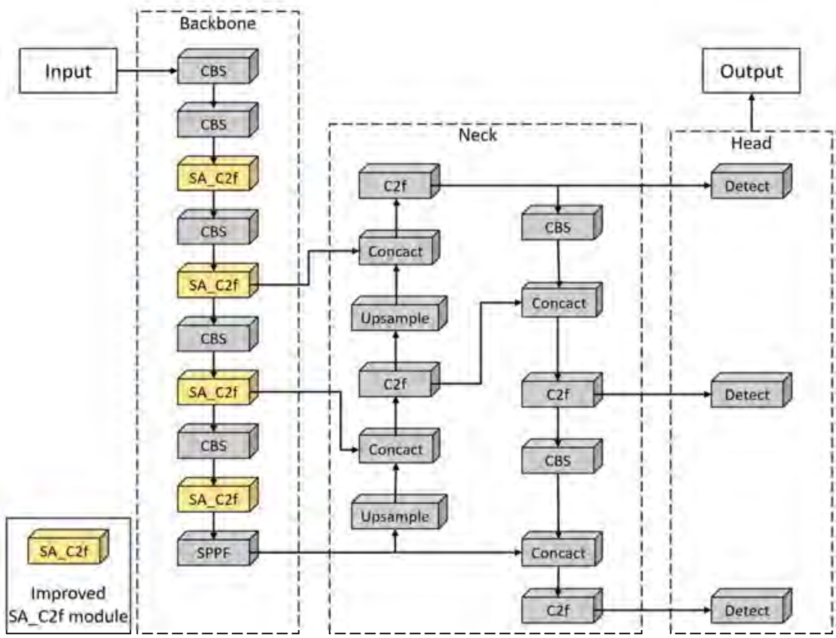
4.1 Dataset construction

Considering the impact of different lighting conditions and board colours on recognition accuracy, this study collected game board images using two methods: first, by automatically generating game boards with different background colours using a Python program, and second, by taking actual photographs of the boards. These images were captured under various lighting conditions, including normal, low, and high brightness. Where the captured images and generated images are approximately 7:3. Specifically, approximately 40% of the captured images are of normal brightness, 35% are of low brightness, and 25% are of high brightness. In addition, the captured images include uneven brightness to enhance the ability of the model to adapt under different lighting conditions. In the generated board images, the background colours cover a wide range of colours such as black, light grey, blue, green and yellow to ensure that the dataset is able to simulate board scenes under different backgrounds. At the same time, some of the images are processed with Poisson noise and Gaussian blur to simulate the possible noise of the board, which further enhances the model adaptability to noise interference. Moreover, different shapes of pieces, such as flat, one-sided convex and notched, are randomly adopted in the captured images, so that the model can adapt to different shapes of pieces. A total of 1,285 game board images were collected and annotated using LabelImg to generate TXT files for training. The constructed dataset is randomly split into training and validation sets in a 7:3 ratio.

4.2 Yolov8 model improvement

The yolov8 model is a detection model in the YOLO series. Compared to previous models, yolov8 offers higher detection speed and accuracy while having fewer parameters. To maintain high detection speed for piece recognition during the game, this study adopts yolov8 as the base model and makes improvements to it.

Figure 6 Improved yolov8 structure (see online version for colours)



The yolov8 model is composed of three key networks: the backbone, neck, and detection head. The multi-level and multi-scale features from the input image extracted by backbone, comprising the CBS, C2f and SPPF modules. Where the C2f module is mainly used for feature extraction and gradient optimisation. The features from the backbone are further integrated and processed by the neck network. A decoupling head structure is adopted by the detection head to complete the final detection.

To ensure high recognition accuracy under different lighting conditions, this study improves the backbone responsible for feature extraction by embedding the SA attention module into the C2f module. This enhances the feature extraction capability of the model, thereby improving overall performance. Figure 6 illustrates the structure of the improved yolov8 model.

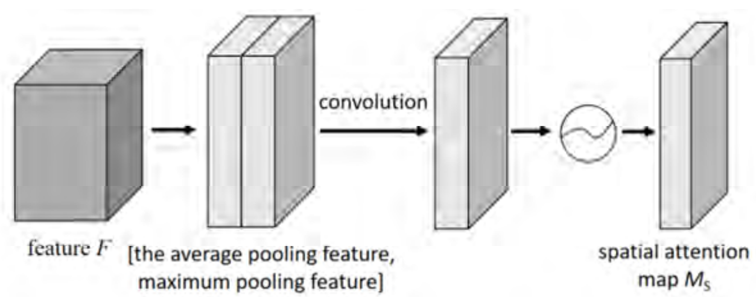
4.2.1 Spatial attention mechanisms

The spatial attention (SA) mechanism serves as an adaptive spatial region selection process (Guo et al., 2022). By mimicking visual system of the human, the SA module makes the model automatically attend to key image regions (Lin et al., 2024). The structure of SA module is shown in Figure 7, and its computational process proceeds as follows: initially, the input feature map F undergoes max-pooling and average-pooling, producing two feature maps of size $H \times W \times 1$. Then these maps are concatenated into a single $H \times W \times 2$ feature map, followed by a 7×7 convolution to generate a single-channel feature map. Finally, the spatial attention map *Ms* is computed using a *Sigmoid* activation function. The computation process is detailed in equation (11).

$$\begin{aligned} Ms(F) &= \sigma\left(f^{7\times 7}([AvgPool(F); MaxPool(F)])\right) \\ &= \sigma\left(f^{7\times 7}([F^s_{avg}; F^s_{\max}])\right) \end{aligned} \tag{11}$$

where F^s_{avg}, $F^s_{\max}$, $f^{7\times 7}$, σ denote the average pooling feature, maximum pooling feature, 7×7 convolution operation, and *Sigmoid* activation function, respectively.

Figure 7 Spatial attention module

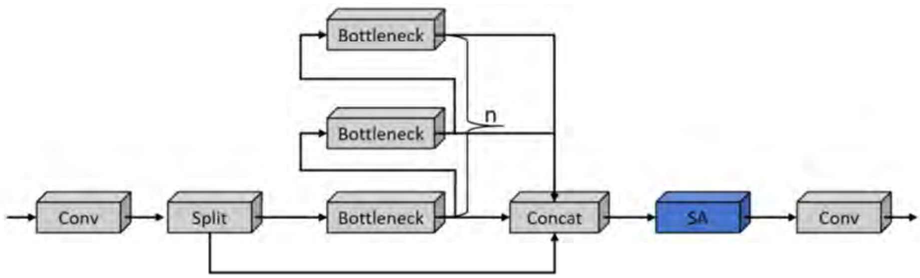


4.2.2 SA_C2f module improvements

In human-machine game scenarios, factors such as lighting variations, background interference, and densely distributed pieces make it challenging for traditional convolution operations to effectively capture the essential features of the pieces, causing a drop in recognition accuracy. This paper proposes an improvement to the C2f module used for feature extraction in the backbone network by embedding the spatial attention (SA) module to resolve this challenge, forming the SA_C2f module. This improvement allows the model to better capture the shape and positional features of the pieces in complex conditions, such as lighting changes, thus improving the accuracy of piece recognition.

By adding the SA module after the concat operation, the model enhances its focus on important regions. Additionally, attention adjustment on the fused feature maps helps handle complex backgrounds and fine details, reducing computational redundancy and improving overall model performance. The improved SA_C2f module's structure is depicted in Figure 8.

Figure 8 SA_C2f module structure (see online version for colours)



4.3 Piece localisation

The piece localisation is crucial to the whole Gomoku game, and the coordinates of the pieces on the board can be accurately recognised to better predict the next move position in the human-machine game. In this paper, for the localisation problem of pieces, on the basis of the division of the corner points of the board and the recognition of pieces, the method of calculating the distance from the centre of the circle to the nearest corner point is introduced, and the calculated distance is compared with the allowed value, when equation (10) is satisfied, it is considered that the piece is located in the corner point, and then the row and column coordinates of the corner point are the row and column coordinates of the piece. Piece localisation is completed by comparing the corner point distance with the allowable value.

$$\begin{aligned} & distance\left[\left(x_{piece}, y_{piece}\right), \left(i_{corner}, j_{corner}\right)\right] \\ & = \sqrt{\left(x_{piece} - i_{corner}\right)^2 + \left(y_{piece} - j_{corner}\right)^2} < allowable\_value \end{aligned} \tag{10}$$

where (x_{piece}, y_{piece}) is the position of the centre of the circle of the piece, (i_{corner}, j_{corner}) is the position of the nearest corner point, and *allowable_value* is the allowable value, which is generally the radius of the piece.

5 Experimental testing

5.1 Experimental platform

The experimental equipment used in this paper is Window 11, the hardware configuration CPU is Intel Core i5-12490F 3.0GHz, 16G RAM, GPU is NVIDIA RTX 4060, 16G RAM; the software environment is Python3.8, Pytorch 2.3.1, CUDA 12.1. The experiments in this paper are divided into three parts, which are: the performance test of the improved model, whether the position of the pieces is placed accurately, and the localisation of the pieces.

5.2 Performance testing of the improved model

In the performance test of SA_C2f-yolov8, both the original and improved models were trained with the following parameters are set as follows: image size of 640 × 640, iteration number of 100, batch size of 8, and the number of working threads is 8.

5.2.1 Evaluation metrics

To assess the effect of the model improvements, precision (*P*), recall (*R*) and mean average precision (*mAP*) are often used as evaluation metrics in target detection (Luo et al., 2024), where precision (*P*) represents the proportion of correct predictions of black pieces(white pieces) to the total number of predictions of black pieces(white pieces) (including incorrect predictions); recall (*R*) represents the proportion of correct predictions of black pieces (white pieces) to the number of actual black pieces (white pieces) on the board. The mean average precision (*mAP*) reflects the model's overall detection capability, with the primary comparison being between *mAP*50 at an IoU threshold of 0.5 and *mAP*50-95 at IoU thresholds ranging from 0.5 to 0.95.

5.2.2 Test results and analysis

In the model performance test of this paper, the improved model SA_C2f-yolov8 is used to conduct experiments on the validation set, which is also judged by combining the recognition of pieces by the model before and after the improvement under different lighting conditions. Table 1 shows the evaluation metrics obtained after the validation of SA_C2f-yolov8 on the validation set, in which the precision rate *P* for both black and white discs is up to 99.9%, the recall rate *R* is 98.6% and 98.5% respectively, the *mAP*50 is 99.4% for both, and the *mAP*50-95 is 96% and 96.2% respectively, which is evident from the above data that the improved model performs well in all the indexes, indicating that it has high precision in recognising piece types.

Table 1 Evaluation metrics of the improved model

Classification	*P (%)*	*R (%)*	*mAP50 (%)*	*mAP50-95 (%)*
Black piece	99.9	98.6	99.4	96
White piece	99.9	98.5	99.4	96.2

Comparing the effect of the model before and after the improvement, Table 2 shows that the improved model outperforms the original in terms of parameters, although the room for improvement is smaller, and SA_C2f-yolov8 is more accurate than yolov8 in the same test environment.

Table 2 Model performance comparison

Model	*P (%)*	*R (%)*	*mAP50 (%)*	*mAP50-95 (%)*
yolov8	99.7	98.9	99.2	96.1
SA_C2f-yolov8	99.9	99.0	99.4	96.4

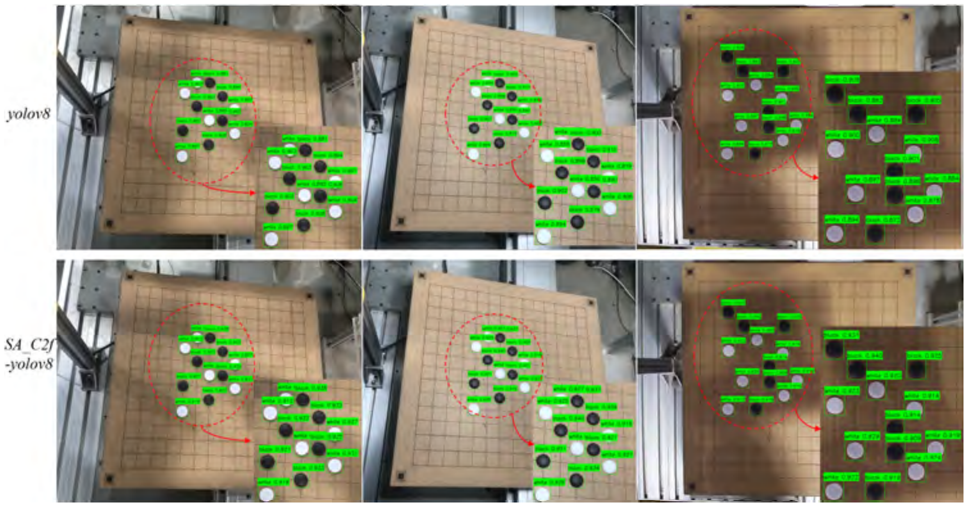
To verify that the improved model meets the requirements of real-time analysis, we conducted tests on dynamic video streams. Since the camera frame rate used in the human-machine gaming device is 30 FPS, the video frame rate used for testing is also set to 30 FPS. The test results show that the average inference time of the improved model is about 13.2 ms, and the corresponding inference frame rate is about 75 frames per second, which is significantly higher than that of 30 FPS, which suggests that, in the case of the camera frame rate of 30 FPS, the model can timely perform the inference for the next frame after each frame processing, and thus can meet the real-time visual analysis requirements of Gomoku.

5.2.3 Visualisation results and analysis

Figure 9 shows the actual recognition effects of the model before and after the improvement, the three sets of photos are affected by different lighting, local occlusion, and tilted external factors of the board. The first and second rows show the recognition performance of the yolov8 and SA_C2f-yolov8 under the same photos, respectively. From these three sets of photos, it can be seen that in the same scene, the confidence level of piece recognition of the improved SA_C2f-yolov8 model basically stays above 0.9, and the confidence level of piece recognition of the original model of yolov8 is roughly in the range of 0.8-0.9, and there are a few of them are above 0.9. Comparing the confidence level of recognition of the two models, from the analysis of the actual recognition effect in Figure 9, the confidence level of the SA_C2f-yolov8 model is higher than that of the yolov8 model in any scenario, which indicates that the effect of the improved model has been improved compared with the original yolov8 model. Meanwhile, in Figure 9(c), the light of the whole board is uneven, and when the light of the region where the pieces are located is poor, the improved model can still detect the pieces normally and correctly determine the types of the pieces, and the confidence level basically stays above 0.9, which fully demonstrates that the improved model can still maintain a higher recognition accuracy under the poor light condition, and can effectively reduce the recognition error caused by the poor light. This also shows that the inclusion of the model in the model can effectively minimise the recognition errors caused by the poor light. This also shows that after adding the attention mechanism to the model, it can effectively extract the target feature area and accurately recognise the type of pieces. From the above analysis, the improved SA_C2f-yolov8 model outperforms the original yolov8 in performance, which is important for the subsequent localisation of pieces.

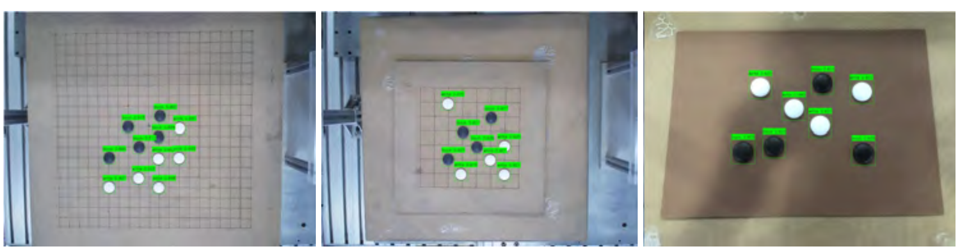
To test the ability of the model to recognise piece types in different scenarios, tests were carried out on boards of different sizes and pure colour paper, and Figure 10 shows the recognition of pieces in different scenarios, respectively. From the figure, it is evident that the overall recognition confidence of pieces stays above 0.9, and the model shows strong adaptability in different scenarios. This may be due to the fact that the function that the model is asked to realise is the judgment of piece types, so only the features of pieces are extracted during the training process instead of features such as the board, and thus the influence of board changes and background interference on the recognition of piece types is small.

Figure 9 Comparison of actual recognition results, (a) recognition of pieces under relatively low brightness (group 1) (b) recognition of pieces under normal brightness (group 2) (c) recognition of pieces under low brightness with shadows (group 3) (see online version for colours)



(a) (b) (c)

Figure 10 Recognition results of piece types in different scenarios, (a) more grid lines on the board (b) less grid lines on the board (c) pure colour paper (see online version for colours)



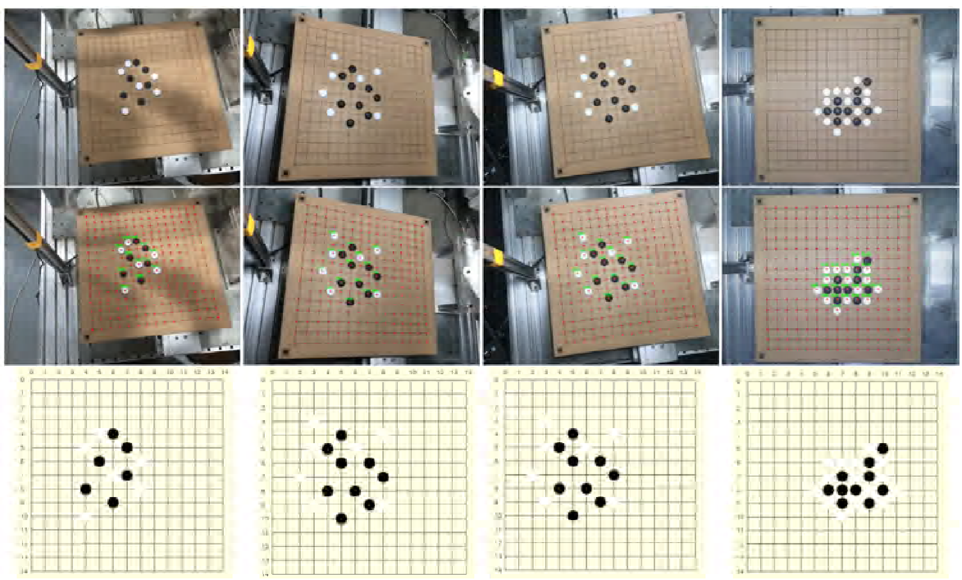
(a) (b) (c)

5.3 Piece placement experiment

Correct placement of pieces means that the pieces are placed on the corner points of the chessboard, and the piece placement experiments in this paper only focus on whether the pieces are correctly placed on the corner points or not. Correct placement of pieces is a prerequisite for accurate localisation. In order to verify the previously mentioned method of calculating the centre-corner distance of pieces, this paper collects several photos of incorrectly placed pieces and sets up the method to display the corresponding coordinates when correctly placed, and to feedback the corresponding signals when incorrectly placed, and to mark the incorrectly placed pieces. Figure 11 shows the case of incorrectly placed pieces, from which it can be seen that the incorrectly placed pieces are marked by red circles and the corresponding labelling is displayed, indicating that the method used for computation mentioned earlier in this paper is applicable to the localisation of pieces in the board.

Figure 11 Incorrect piece placement (see online version for colours)

Figure 12 Piece localisation diagram, (a) the board is rotated in normal brightness with localised shadows (group 1) (b) the board is tilted (group 2) (c) the board is rotated in stronger brightness (group 3) (d) the board is in lower brightness (group 4) (see online version for colours)



(a) (b) (c) (d)

5.4 Piece localisation experiment

The main factors affecting the localisation of pieces are: different brightness, local shadows, and the tilt and rotation of the board. In this paper, we carry out chess localisation experiments under the influence of the above influencing factors and verify the accuracy of chess localisation by combining the virtual board. Figure 12 shows four groups of photos of chess games under different factors, the first group of which is the local shadows under normal luminance and the board is in the rotating state; the second group of the board is in the tilting state; the third group of the photos is in the case of stronger luminance and the board is in the rotating state; and the fourth group of the photos is in the case of lower luminance. As shown in Figure 12, the first, second and third rows are the original image, the board image by processed and the virtual board, respectively. The photos of games with different influences have accurate delineation of the corner points of the board, it can be seen that the types and positions of the pieces correspond to each other one by one by comparing the positions of the pieces on the physical board with those on the virtual board, which indicates that the method of piece localisation in this paper is accurate, and it can adapt to the factors of different luminance, tilt, and rotation.

6 Conclusions

Aiming at the requirements of move strategy and control drive of human-machine gaming equipment on the accuracy of the position of chess pieces, as well as the problems of uneven brightness, low brightness, camera tilting and other complexities in the process of gaming that lead to the low accuracy of recognition and localisation of the position of Gomoku pieces, this paper presents a method of recognising and localising the Gomoku game based on image processing and deep learning. For the problem of obtaining the corner points of the board, extract the region of the board as well as the location of the vertices by image processing technology, and correct them according to the principle of perspective transformation, and obtain the corner points of the board by interpolation calculation method. To improve the accuracy of piece recognition in complex situations such as different brightness, the SA attention mechanism is introduced into the C2f module of the yolov8 model. For the piece recognition problem, the SA attention mechanism is embedded in the C2f module of the yolov8 model and the model is trained by collecting different images, and then the trained model is used to recognise the Gomoku images and obtain the category information of the Gomoku pieces. Finally, the recognised pieces are compared with the corresponding corner points to determine the position of the pieces of different categories, so as to realise the accurate identification and localisation of the pieces in the Gomoku game.

To test the effectiveness of the proposed method in this paper, the SA_C2f-yolov8 model is validated on the validation set, and the recognition and localisation accuracy of the pieces are examined on the images of Gomoku games with different situations such as uneven brightness and low brightness. The experimental results show:

1 The accuracy *P*, recall *R*, *mAP*50, and *mAP*50-95 obtained from the SA_C2f yolov8 model validation are 99.9%, 99.0%, 99.4% and 96.4%, respectively. The confidence level is basically more than 90% in the actual recognition, and the overall performance is improved compared with the yolov8 model.

2 The improvement of Gomoku piece recognition and localisation accuracy is obvious in complex situations such as uneven brightness, low brightness and camera tilt, which can well satisfy the functions of different Gomoku piece recognition and localisation during Gomoku games.

Although the method proposed in this paper achieves good results in the recognition and localisation, the recognition and localisation accuracy of Gomoku pieces may be affected under extreme lighting conditions such as very low light. In addition, the effectiveness of the trained model is limited on low configuration devices or under resource constraints. Future research will focus on optimising the recognition and localisation of pieces in extreme environments and exploring optimised solutions for resource-constrained devices so that the models can be readily deployed on low-configuration devices.

Declarations

This research was funded by the Science Rearch Fund Project of Yunnan Provincial Department of Education (2024Y611).

References

Chen, L., Wang, Z.Y., Huang, J.L. et al. (2020) 'Design of artificial intelligence experiment platform based on Chinese chess robot', *Radio Engineering*, Vol. 50, No. 10, pp.839–842.

Fabris, G., Scalera, L. and Gasparetto, A. (2024) 'Playing checkers with an intelligent and collaborative robotic system', *Robotics*, Vol. 13, No. 1, p.4.

Guo, M., Xu, T., Liu, J. et al. (2022) 'Attention mechanisms in computer vision: a survey', *Computational Visual Media*, Vol. 8, No. 3, pp.331–368.

Larregay, G., Pinna, F., Avila, L. et al. (2018) 'Design and implementation of a computer vision system for an autonomous chess-playing robot', *Journal of Computer Science and Technology*, Vol. 8, No. 1, pp.1–11.

Li, G.Y. and Wang, T. (2020) 'Research on concrete crack detection based on morphological processing and feature analysis', *Building Structure*, Vol. 50, No. S2, pp.529–533.

Li, X., Su, J.N., Zhang, Z.P. et al. (2024) 'Recognition method for fine-grained product styles based on deep learning', *Computer Integrated Manufacturing Systems*, Vol. 30, No. 3, pp.1011–1022.

Lin, C.X., Hu, X.J., Zhan, Y.F. and Hao, X.T. (2024) 'MobileNetV2 with spatial attention module for traffic congestion recognition in surveillance images', *Expert Systems with Applications*, Vol. 255, Part D, p.124701.

Luo, G.F., Wang, Y., Li, H. et al. (2024) 'Unsafe behavior real-time detection method of intelligent workshop workers based on improved yolov5s', *Computer Integrated Manufacturing Systems*, Vol. 30, No. 5, pp.1610–1619.

Mao, L.M., Zhu, P.Y., Lu, Z.L. et al. (2017) 'Chessboard recognition for Gomoku robot based on LabVIEW', *Computer Engineering and Design*, Vol. 38, No. 1, pp.242–246.

Mishra, B., Gopal, K.M., Patnaik, S. and Paikaray, B.K. (2023) 'Identification and detection of brain tumour using deep learning-based classification MRI applied using neural network and machine learning algorithm', *International Journal of Reasoning-based Intelligent Systems*, Vol. 15, Nos. 3–4, pp.228–234, DOI: 10.1504/IJRIS.2023.136350.

Ou, J.C. (2020) *Research on the Game Platform of Gomoku Man-Machine Based on DOBOT Manipulator*, Shanghai University of Engineering Science, Shanghai.

Patil, P., Deokate, S.T., Bhoite, A. et al. (2023) 'GAN-enhanced medical image synthesis: augmenting CXR data for disease diagnosis and improving deep learning performance', *Journal of Electrical Systems*, Vol. 19, No. 3, https://doi.org/ 10.52783/jes.651.

Prusty, S., Dash, S.K. and Patnaik, S. (2022) 'A novel transfer learning technique for detecting breast cancer mammograms using VGG16 bottleneck feature', *ECS Transactions*, Vol. 107, No. 1, p.733, https://doi.org/10.1149/10701.0733ecst.

Prusty, S., Patnaik, S., Dash, S.K. et al. (2024a) 'SEMeL-LR: an improvised modeling approach using a meta-learning algorithm to classify breast cancer', *Engineering Applications of Artificial Intelligence*, Vol. 129, p.107630, https://doi.org/ 10.1016/j.engappai.2023.107630.

Prusty, S., Patnaik, S., Dash, S.K., Prusty, S.G.P., Rautaray, J. and Sahoo, G. (2024b) 'Predicting cervical cancer risk probabilities using advanced H20 AutoML and local interpretable model-agnostic explanation techniques', *PeerJ Computer Science*, Vol. 10, p.e1916, https://doi.org/10.7717/peerj-cs.1916.

Silver, D., Huang, A., Maddison, C.J. et al. (2016) 'Mastering the game of Go with deep neural networks and tree search', *Nature*, Vol. 529, No. 7587, pp.484–489.

Silver, D., Schrittwieserh, J., Simonyan, K. et al. (2017) 'Mastering the game of Go without human knowledge', *Nature*, Vol. 550, No. 7676, pp.354–359.

Wang, Y.J., Zhang, Y.B., Wu, Y.Y. et al. (2020) 'General chess piece localization method under uneven illumination', *Journal of Computer Applications*, Vol. 40, No. 12, pp.3490–3498.

Wang, Z.A., Wang, H.W. and Zang, Y.L. (2021) 'A replay method for Gomoku chessboard based on machine vision', *2021 Asia-Pacific Conference on Communications Technology and Computer Science (ACCTCS)*, IEEE, Piscataway, pp.266–270.

Wu, Y.Q., Chen, D.L., Wang, Y. et al. (2019) 'Intelligent man-machine Gomoku game robot design', *Chinese Journal of Electron Devices*, Vol. 42, No. 4, pp.968–972.

Xue, R.Q., Yuan, Q.S. and Su, J.D. (2023) 'Design and implementation of man-machine playing device for Gomoku game', *2023 5th International Conference on Frontiers Technology of Information and Computer (ICFTIC)*, IEEE, Piscataway, pp.673–678.

Yuan, H. and Li, W. (2022) 'Machine learning-based multi-target detection method for the complex mine environment', *International Journal of Reasoning-based Intelligent Systems*, Vol. 14, No. 4, pp.201–207, DOI: 10.1504/IJRIS.2022.126654.

Zhang, X. (2022) *System Design and Key Technology Research of Chess-playing Robot*, Southwest Jiaotong University.

Zhang, Y. and Gu, F. (2022) 'A Compact and Low-cost Gomoku Robot', *2022 International Conference on Service Robotics (ICoSR)*, IEEE, Piscataway, pp.62–66.

Zhao, X. (2022) *Research on the Design and Visual Algorithm of Go Robots*, Zhejiang Sci-Tech University, Zhejiang.

Zhu, R., Qi, Y. and Zhang, Y. (2024) 'Lightweight PCB defect detection algorithm incorporating partial convolution and parameter-free attention mechanisms', *Computer Integrated Manufacturing Systems*, 9 September, pp.1–18.