Robotica (2026), **44**, pp. 1703–1720
doi:10.1017/S0263574726103695

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Hole search strategy for robotic peg-in-hole assembly under initial positioning uncertainty

Xianda Wang[1,2,3], Qinghua Lu[2,3], Weilin Chen[2,3], Shijie Tang[1,2,3], Qinghua Zhang[2,3], Lufeng Luo[2,3] and Mingyou Chen[2,3]

[1]College of Mechanical and Electric Engineering, Changchun University of Science and Technology, Changchun, China
[2]School of Mechatronics Engineering and Automation, Foshan University, Foshan, China
[3]Guangdong Provincial Key Laboratory of Industrial Intelligent Inspection Technology, Foshan University, Foshan, China
Corresponding authors: Qinghua Lu; Email: qhlu@fosu.edu.cn; Weilin Chen; Email: weilin.chen@fosu.edu.cn

Received: 11 December 2025 UTC; **Revised:** 21 April 2026 UTC; **Accepted:** 15 May 2026 UTC;
First published online: 11 August 2026 UTC

Keywords: peg-in-hole assembly; hole search strategy; contact state; trajectory optimization; convergence analysis

Abstract

Blind search methods are widely used in the search phase of robotic peg-in-hole assembly. However, these methods are sensitive to the initial contact position and are time-consuming. Additionally, in scenarios with a limited search space, such as thin-walled holes, the hole search process is prone to failure due to loss of contact between the peg and the hole. To address this issue, this paper proposes a hole search strategy that uses visual detection of the contact state between the peg and the hole and peg position adjustment based on the detection results. First, a contact-state classification model under fixed single-view conditions is constructed, converting the disordered relative positional relationships between the peg and the hole into discrete contact states, laying a theoretical foundation for the vision-guided hole search process. Second, search trajectories are generated based on the contact states, and trajectory optimization is further performed through visual error analysis, thereby overcoming hole search failures caused by visual occlusion and camera mounting deviations. Finally, convergence analysis shows that the peg slides into the hole when the classification accuracy is greater than 50%. Robotic assembly experiments further demonstrate the effectiveness of the proposed contact-state classification model and search strategy in tasks such as thin-walled hole insertion and USB Type-A(USB-A) connector insertion.

1. Introduction

Robots are widely used in industrial manufacturing owing to their high flexibility and level of automation, and robotic assembly represents one of their most important applications. Among various assembly tasks, peg-in-hole assembly is a basic task that appears extensively in mechanical manufacturing, automotive production, aerospace, and consumer-electronics industries [1–3]. Research on peg-in-hole assembly typically focuses on the search phase and the insertion phase. The search phase is particularly challenging because it is affected by sensor noise, positioning errors, and environmental uncertainties, and it is also a critical factor that determines the success of the assembly task.

To enhance a robots' ability to perceive its environment, vision, force, and multi-sensor fusion technologies have been applied to the hole search task. Among these, visual technology is widely used due to its ability to acquire environmental information without physical contact. Puang et al. [4] employed a keypoint extraction network to learn keypoint representations from images, and a visual servoing network predicted the robot's motion based on these keypoints, achieving an assembly with a peg-hole clearance of 4 mm. Sileo et al. [5] perform point cloud matching between the 3D reconstruction model and the CAD model to determine the pose of the hole, and further improves the position estimation accuracy using a convolutional neural network. Finally, a blind search strategy is integrated to accomplish the hole-searching task. Triyonoputro et al. [6] used learning-based visual servoing to quickly

© The Author(s), 2026. Published by Cambridge University Press.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 22:56:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0263574726103695

move the peg closer to the hole, followed by a spiral search to achieve alignment between the peg and the hole. Wang et al. [7, 8] proposed a visual admittance control method based on image feature space for peg-in-hole assembly, which achieves accurate positioning through visual servoing. By introducing adaptive stiffness, force/vision fusion, and asymmetric Lyapunov constraints, the approach enhanced the system's environmental adaptability. Wu et al. [9] established a two-dimensional hole-searching model by fusing signals of vision detection and fuzzy force as conditions for state transitions, and employed a semi-supervised learning network to optimize the hole-searching process.

To address the issues of visual sensors being susceptible to occlusion and having limited accuracy in practical applications, blind search methods that do not rely on visual information have been widely studied and applied. The advantages of these methods lie in their simplicity and strong applicability. Blind search methods include multiple search trajectories, such as spiral trajectory [10–12], random trajectory [13], and Lissajou trajectory [5, 14], among others. However, blind search methods usually require good initial positioning and traverse long paths and times [15–17]. Park et al. [11] reduced the search range by actively tilting and translating the peg in a specific direction, thereby reducing the randomness of global blind search, and improved the spiral force trajectory to a partial spiral force trajectory, further reducing the uncertainty of search time. Kang et al. [12] determined the hole search range based on the visually estimated hole pose and system configuration, and completed the hole searching using a spiral search. Liu et al. [18] proposed an assembly method based on the geometric features of the constraint region, in which a multi-step sliding strategy is used to eliminate positional errors and complete the assembly. Kim et al. [19] proposed a shape recognition and hole detection algorithm based on force, and the direction of the hole was obtained according to the hole detection algorithm. Wu et al. [20] adjusted the peg's pose based on feedback from a six-axis force/torque sensor and achieved alignment between the workpiece and the hole through vertical insertion, force and position threshold evaluation, and iterative fine-tuning. Tang et al. [21] estimated the pose deviation based on the force and geometric analyses of the three-point contact state between the peg and the hole and completed the peg-hole alignment before insertion. However, assembly methods based on contact information typically rely on stable and continuous contact, which not only increases control complexity but also reduces assembly efficiency.

With the increasing complexity of robotic manipulation tasks, reinforcement learning [22–24] and imitation learning [25, 26] have been introduced into robotic manipulation tasks because they can automatically acquire manipulation policies from experience. However, such methods usually face problems such as high training costs and low sample efficiency and are difficult to meet the stability and reliability requirements of industrial production [15, 27, 28]. In addition, these technologies require a high level of skill from practitioners, which limits their application in practice.

In summary, despite extensive research on hole search strategies, this issue remains a significant challenge. To address these issues, this paper proposes a vision-guided search strategy to perform a hole search in robotic peg-in-hole assembly. The core idea of the proposed strategy is to continuously guide the peg in adjusting its position based on relative position information, without explicitly estimating the positional error between the two. The main contributions of this paper are as follows:

- A contact state classification model based on relative positional relationships between the peg and the hole is proposed, which discretizes the continuous and disordered relative positional relationships into a finite number of contact states. Based on this, contact state detection is achieved from a fixed single-viewpoint, providing theoretical support for a vision-guided search process that does not require precise position estimation.
- The search trajectories are generated based on the contact states between the peg and hole, and the trajectory is optimized according to the analysis results of visual detection errors, thereby addressing the impact of visual occlusion and camera mounting deviations on the search phase.
- This paper applies a biased random walk model to the convergence analysis of the hole search strategy, proving that even with classification errors, as long as the classification accuracy is greater than 50%, the peg can still slide into the hole with a fixed step size.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 22:56:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0263574726103695

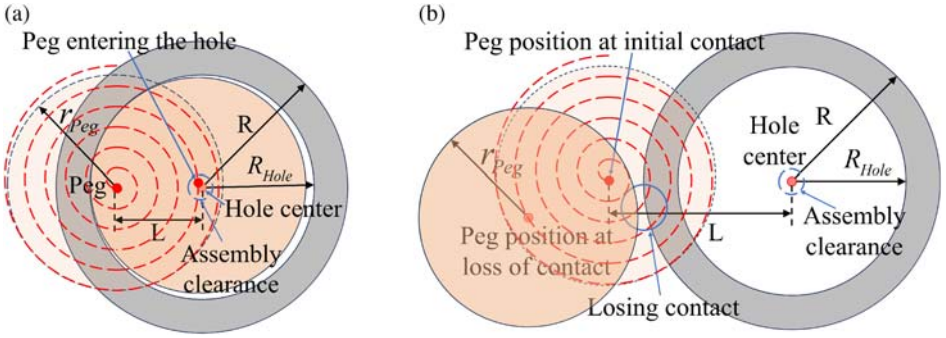


Figure 1. *Applicability of spiral search in the thin-wall hole search phase. (a) When the initial contact condition is satisfied $L \leq (R + R_{Hole})/2$, the spiral trajectory can complete the hole search task. (b) When the condition is satisfied $R + r_{peg} > L > (R + R_{Hole})/2$, the hole search task fails.*

The remainder of this paper is organized as follows. Section 2 introduces the background and challenges of the hole search task, proposes a corresponding search strategy. Section 3 constructs a peg–hole contact-state classification model, optimizes the search trajectory, and proves the convergence of the strategy. Section 4 verifies the effectiveness and robustness of the proposed strategy through experiments. Section 5 summarizes the contents of this paper.

2. Hole search strategy

2.1. Problem statement

In the planar hole search task, even when the positioning error is relatively large, the hole search task can still be completed by using time-consuming blind search methods. However, such methods are not suitable for thin-walled holes. In this study, the initial positioning error is assumed to be smaller than the sum of the peg radius and the outer radius of the thin-walled hole. Under this assumption, the peg may lose contact with the thin-walled hole during the hole search process, leading to task failure [17]. The essence of the hole search process lies in guiding the peg into the assembly clearance region formed by the dimensional difference between the peg and the hole, which is typically small. Therefore, the key to a search strategy is to ensure that the search trajectory can cover this region while maintaining contact constraints. As shown in Figure 1(a), when the initial contact position satisfies the condition $L \leq (R + R_{\text{Hole}})/2$, the spiral search can complete the hole search task, whereas using other trajectories may risk losing contact between the peg and the hole, leading to failure of the hole search task. On the other hand, as shown in Figure 1(b), when the initial contact position satisfies the condition $R + r_{\text{peg}} > L > (R + R_{\text{Hole}})/2$, the effective contact area between the peg and the thin-walled hole is small, and the hole will lose contact during the blind search process, resulting in the failure of the hole search.

2.2. Hole search strategy

When humans perform peg-in-hole assembly tasks, they typically rely on visual observation of the relative positional relationship between the peg and the hole for position adjustment, and combine tactile information to complete the assembly. For robots, visual information can be acquired via cameras, and contact forces can be measured using force sensors. However, visual cues are often affected by occlusion, resulting in incomplete feature representation. Moreover, the lack of texture on the peg and the hole surfaces further increases the difficulty of feature extraction.

Inspired by human assembly behavior, this study develops a vision-guided hole search strategy, and the overall framework is illustrated in Figure 2. First, relative position features of the peg-hole contact are extracted using vision. Then, the robot's movement direction is adjusted based on the visual

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 22:56:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0263574726103695

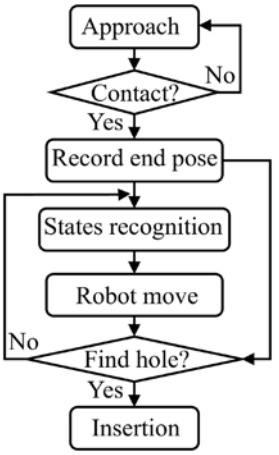


Figure 2. *Hole search strategy flow chart.*

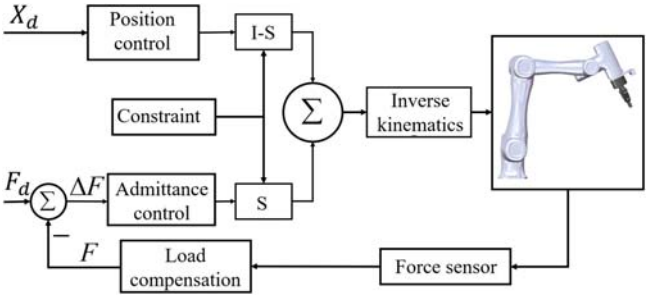


Figure 3. *Hybrid force/position control.*

detection results. The assembly process consists of three phases: approach, search, and insertion; however, this paper primarily focuses on the search phase. To ensure the accurate and orderly completion of tasks at each assembly phase, this paper designs specific indicators to signal the conclusion of each phase, including the robot's position information and contact forces/torques. When the peg first contacts the hole and enters the search phase, the system records the position information at this time as an important basis for subsequent phase judgments. During the search phase, the robot continuously adjusts the peg position until the insertion depth exceeds the threshold, at which point it enters the insertion phase.

2.3. Control frameworks

In assembly tasks, force control technology is required to achieve stable contact between the robot and the environment. Admittance control is the most widely used force control method, in which the robot system is typically modeled as a second-order mass-spring-damper system, as shown in Eq. (1).

$$\begin{cases} \boldsymbol{M}_d\ddot{\boldsymbol{x}}_e + \boldsymbol{B}_d\dot{\boldsymbol{x}}_e + \boldsymbol{K}_d\boldsymbol{x}_e = \boldsymbol{F} - \boldsymbol{F}_d \\ \boldsymbol{x}_e = \boldsymbol{x} - \boldsymbol{x}_d \end{cases} \tag{1}$$

where M_d, B_d, K_d are the positive-definite virtual inertia, damping, and stiffness matrices, respectively; x is the current pose, x_d is the desired pose; F is the current contact force/torque; and F_d is the desired contact force/torque.

As shown in Figure 3, the control system combines position control and force control to construct a robot hybrid force/position control framework. The control mode is switched through the 6×6 diagonal matrix S, when the diagonal element of the matrix is 1, it means force control, and 0 means position control.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 22:56:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0263574726103695

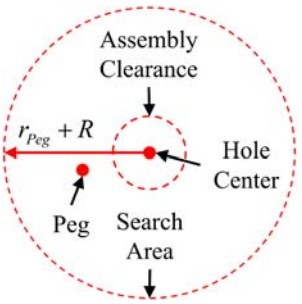


Figure 4. *Hole search model.*

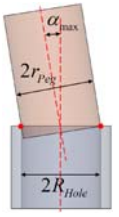


Figure 5. *Critical state of two-point contact inside a hole.*

The pose increment under hybrid force/position control can be expressed as:

$$\Delta \boldsymbol{p} = (\boldsymbol{I} - \boldsymbol{S}) \cdot \boldsymbol{\Phi}(\boldsymbol{X}_d) + \boldsymbol{S} \cdot \Gamma(\Delta \boldsymbol{F}) \tag{2}$$

Where Δp is the position increment of the robot's end effector, Φ represents the position control method, Γ represents the force control method, and ΔF is the force/torque error.

3. Methodology

3.1. Assembly analysis

Figure 4 shows the hole search model of the proposed strategy. Due to the dimensional difference between the peg and the hole, an assembly clearance region is formed. When the peg enters the assembly clearance region, it will slide into the hole due to the loss of support force. However, the initial positional error between the peg and the hole usually lies outside this region. Therefore, the goal of the search phase is to guide the peg into the assembly clearance region.

In circular peg-in-hole assembly, due to pose deviation between the peg and the hole, two-point contact inside the hole is a common contact form. As shown in Figure 5, under the two-point contact state inside the hole, the peg is in the critical geometric condition where it can just enter the hole, corresponding to the maximum allowable tilt angle between the peg and the hole. When the relative tilt angle between the peg and the hole exceeds this critical angle, the peg can no longer enter the hole due to geometric constraints. Based on the geometric analysis of this critical state, under the two-point contact condition inside the hole, the maximum relative inclination angle between the peg and the hole is given by $\alpha_{\max} = \pm \arccos\left(r_{\text{peg}}/R_{\text{Hole}}\right)$.

3.2. Coordinate system conversion

As shown in Figure 6, there are multiple coordinate systems involved in the assembly task: robot base coordinate system {B}, camera coordinate system {C}, and robot tool coordinate system {T}. The origin of {C} is located at the optical center of the RGB camera, and the origin of {T} is located at the center of the peg end face.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 22:56:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0263574726103695

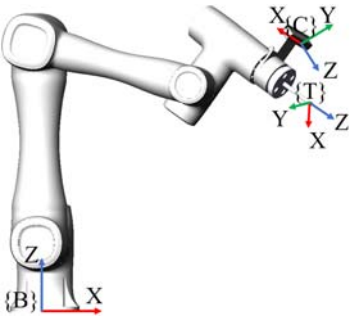


Figure 6. *Coordinate systems in assembly systems.*

In the assembly process, both force and motion are defined in the coordinate system {T}, while motion control is performed in the coordinate system {B}. Therefore, the pose and force/torque conversion must be performed between the two coordinate systems to achieve consistency between the perception information and the motion. Formula (3) represents the translation and rotation in the tool coordinate system, and the mapping to the base coordinate system through coordinate transformation. Formula (4) represents the force/torque transformation relationship between different coordinate systems, aiming to convert the force/torque driving the robot's motion from coordinate system {T} to {B}, in order to achieve motion regulation in admittance control.

$$\begin{cases} \mathbf{T} = \begin{bmatrix} \mathbf{R} & \mathbf{p} \\ 0 & 1 \end{bmatrix} \\ {}^{\text{base}}_{\text{new}}\mathbf{T} = {}^{\text{base}}_{\text{ee}}\mathbf{T} \cdot {}^{\text{ee}}_{\text{new}}\mathbf{T} \end{cases} \tag{3}$$

$$\begin{cases} \mathbf{D} = \mathcal{S}\left({}^{B}\mathbf{p}_{O}\right) = \begin{bmatrix} 0 & -p_z & p_y \\ p_z & 0 & -p_x \\ -p_y & p_x & 0 \end{bmatrix} \\ \begin{bmatrix} {}^{M}\mathbf{F} \\ {}^{M}\mathbf{M} \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{DR} & \mathbf{R} \end{bmatrix} \begin{bmatrix} {}^{N}\mathbf{F} \\ {}^{N}\mathbf{M} \end{bmatrix} \end{cases} \tag{4}$$

where p is the position vector, R is the rotation matrix, ${}^{\text{ee}}_{\text{new}}T$ represents the homogeneous transformation matrix that performs pose transformation in the tool coordinate system, ${}^{\text{base}}_{\text{ee}}T$ represents the homogeneous transformation matrix of the tool coordinate system relative to the base coordinate system, D denotes the skew-symmetric matrix of the position vector, and ${}^{M}\mathbf{F}$, ${}^{N}\mathbf{F}$, ${}^{M}\mathbf{M}$, ${}^{N}\mathbf{M}$ represent forces and torques in different coordinate systems.

3.3. Elementary movement

Robotic motion control is typically categorized into position control and force control. In practical applications, robots typically employ the position control mode, which offers the advantages of high position-tracking accuracy and a relatively mature implementation framework. Under the force control mode, the robot is driven to adjust its position based on a set desired force. Since the underlying architecture of most commercial robots is based on position control, they typically employ admittance control to convert external force information into position-correction values, thereby indirectly achieving force control functionality. Because the adjustment is performed along the surface of the hole, both the desired force and the position adjustments are defined with respect to the coordinate system {T}.

The robot's action can be defined as a desired force vector or position vector:

$$\mathbf{a}_f = [F_x \quad F_y \quad F_z] \tag{5}$$

$$\mathbf{a}_p = [\Delta x \quad \Delta y \quad \Delta z] \tag{6}$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 22:56:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0263574726103695

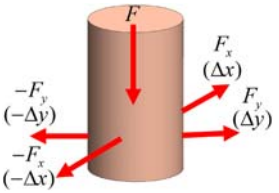


Figure 7. *Elementary movement.*

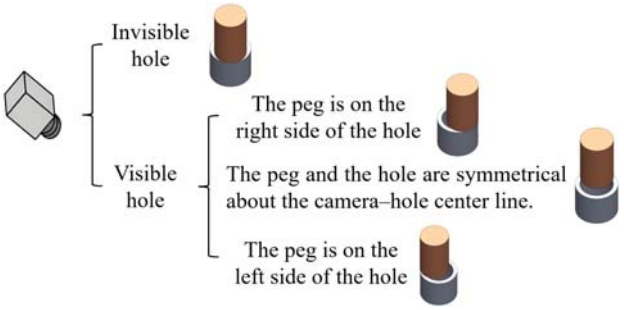


Figure 8. *Contact state classification model under fixed single-view conditions.*

where F_x, F_y, F_z is the desired force along the XYZ direction of the coordinate system {T}, Δx, Δy, Δz are the motion steps along the XYZ direction of the coordinate system {T}. The elementary movement formed by each component of the action vector is shown in Figure 7. The robot's action can be defined as a combination of one or more elementary movements.

3.4. Contact state model

The core of visual perception is not to directly quantify the positional error between the peg and the hole through the visual system, but to perform contact state classification based on the positional relationships and their appearance in the camera image. Since there is no explicit category definition for the relative positional relationships between the peg and the hole, it is necessary to manually establish classification rules. These rules transform the unordered relative positional relationships into an ordered and discrete set of contact states, facilitates the neural network in learning the visual features associated with different contact states. As shown in Figure 8, when contact occurs between the peg and the hole, under a fixed single-view condition, the contact state can be classified based on the visibility of the hole in the camera's field of view and the relative positional relationships between the peg and the hole. This observation provides a basis for constructing the subsequent classification rules.

Based on the above analysis, we classify the contact states between the peg and the hole into four categories and establish a plane coordinate system{O} with the center of the hole's upper surface as the origin to describe the specific characteristics of each contact state, as shown in Figure 9. The camera adopts the eye-in-hand configuration, with the assumption that the projection of the RGB camera's mounting position onto the XOY plane lies on the positive X-axis, while the camera is oriented toward the center of the hole. Under this assumption, the specific classification rules are as follows:

1. When the center point of the peg end face is along the positive direction of the X-axis, and the hole cannot be seen from the camera's perspective, it is defined as the front class;
2. When the center point of the peg end face is along the negative direction of the X-axis, and both the peg and the hole are symmetrical about the X-axis, it is defined as the back class;
3. In the camera perspective, the peg is located on the left side of the hole, which is defined as the left class;

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 22:56:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0263574726103695

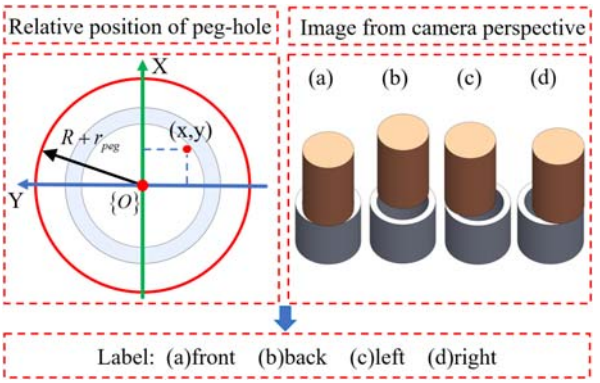


Figure 9. *Dataset acquisition and contact state division.*

4. In the camera perspective, the peg is located on the right side of the hole, which is defined as the right class.

The above classification rules are also the foundation for establishing contact datasets in the future, the center of the hole is taken as the origin, and the circle with a radius of $R + \mathbf{r}_{\text{peg}}$ is taken as the acquisition area of the contact state images, as shown in Figure 9. When the peg and hole are in contact, the camera acquires as many images as possible to construct the dataset. Both the acquisition of the dataset and the detection of the contact states rely on visual information. However, due to the discrepancy between the actual camera mounting pose and the ideal mounting pose, it is impossible to directly define image labels strictly according to the above classification rules. Therefore, after acquiring the images, the actual display effect in the image needs to be considered when labeling the data.

3.5. Search trajectory generation and optimization

This paper defines the contact state recognition between the peg and the hole as an image classification problem. Since the region of interest typically occupies only a small portion of the image, conventional full-image classification methods struggle to focus on relevant features effectively. Therefore, the YOLOv8 model is adopted to enable both localization and recognition of the contact state. The model structure adopts a lighter and more efficient C2f module, combined with improvements such as decoupling head, which significantly improves detection accuracy while ensuring real-time performance. With their excellent performance, YOLOv8 and its improved versions have been widely used in various object detection tasks [29, 30]. Given the clear target structure and low background complexity in this research task, The YOLOv8n model, as the most lightweight network architecture, is sufficient to meet the requirements.

In this study, a camera acquires image sequences in real time and inputs them frame by frame into a pre-trained YOLOv8n model to detect the current contact state between the peg and the hole. Based on the detection results, corresponding position adjustment actions are generated to guide the peg to move in a direction that reduces the positional error between the peg and the hole. A mapping relationship between the contact states and the position adjustment actions is established, as shown in Figure 10.

The contact state detection results depend not only on the relative positional relationship between the peg and the hole but are also influenced by the display conditions from the camera's viewpoint. Therefore, camera mounting angle deviation can affect visual detection results and further influence the search phase. The adjustment motion of the peg is designed based on the coordinate system {T}, while visual detection is performed in the camera coordinate system {C}. When the projection of the camera coordinate system {C} onto the XOY plane of coordinate system {T} is not aligned with the coordinate axes, the visual detection results will be affected. If there is an angle deviation α between the expected and actual camera mounting positions, some of the left- or right-class objects under the

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 22:56:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0263574726103695

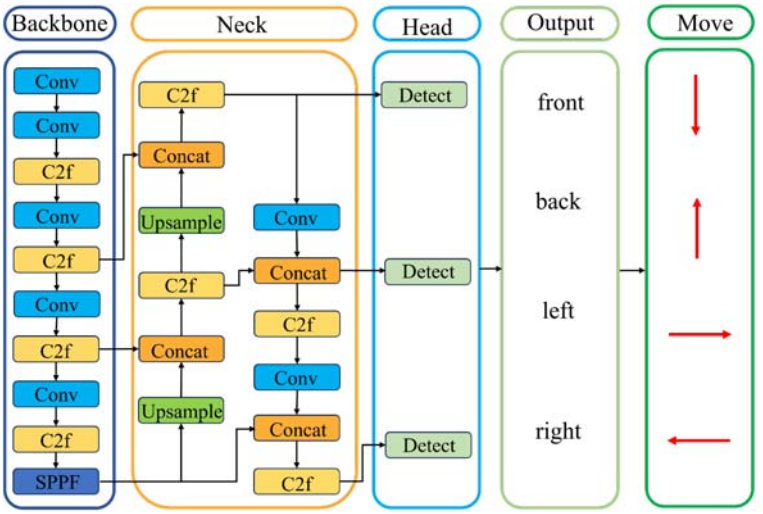


Figure 10. *YOLOv8n-based contact state detection and position adjustment.*

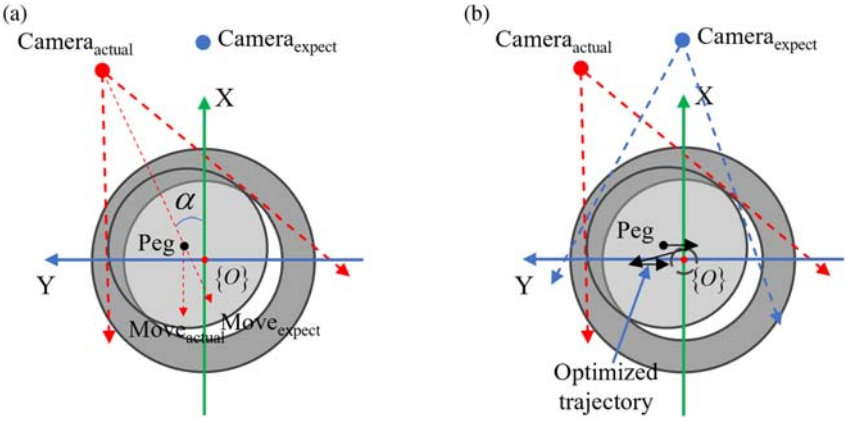


Figure 11. *(a) Motion error. (b) search trajectory optimization.*

ideal camera pose may be recognized as front class in the actual mounting pose. In conventional circular peg-in-hole assembly tasks, variations in the camera mounting angle can affect the classification results, causing a discrepancy between the expected and actual motion. As shown in Figure 11(a), the expected motion of the peg is $Move_{\text{expect}}$, while the actual motion is $Move_{\text{actual}}$. According to the properties of the cosine function, when $\alpha < 45°$, the main direction of peg adjustment remains along the desired adjustment direction, as shown in formula (7). Therefore, the effect of camera mounting angle deviation on conventional circular peg-in-hole assembly is limited.

$$\begin{cases} Move_{\text{actual}} \cdot \cos\alpha < Move_{\text{expect}} \\ Move_{\text{actual}} \cdot \cos\alpha > Move_{\text{actual}} \cdot \sin\alpha, \quad \alpha < 45° \end{cases} \tag{7}$$

Camera mounting angle deviation has little impact on conventional circular peg-in-hole assembly, and hole search can be successfully performed without any specific compensation. However, in assembly tasks with visual occlusion, relying solely on contact-state-based single-direction position adjustments is often insufficient to ensure a reliable search process. In such cases, displacements in other directions can be superimposed while adjusting the peg position to generate a zigzag or other search trajectory, as shown in Figure 11(b), in order to expand the area scanned by the search path and allow the peg to enter the previously difficult to reach area, effectively solving the problem of visual obstruction in

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 22:56:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0263574726103695

the search process. For a rectangular peg-in-hole assembly task, even if there are no camera mounting angle deviations, the front side of the rectangular peg can still occlude the left and right sides when the peg and hole are in close proximity, thereby affecting the detection of contact states on both sides. In addition, visual accuracy is limited, and it is also necessary to compensate for the lack of visual accuracy by optimizing the search trajectory.

3.6. Proof of convergence of hole search strategy

3.6.1. Problem modeling

Let the positional error between the peg and the hole be defined as a two-dimensional vector $\mathbf{d}_k = (x_k, y_k)$, where k represents the k-th adjustment step. The vision-based classification module divides the current contact state into four classes (e.g. front, back, left, and right). According to the classification result, the robot executes a predefined motion with a fixed step length $h > 0$. The control policy can be expressed as:

$$\begin{cases} d_{k+1} = d_k + \Delta d_k \\ \Delta d_k = \begin{cases} -h\dfrac{d_k}{\|d_k\|} & \text{correct probability } p \\ +hu_s & \text{Incorrect probability } 1-p \end{cases} \end{cases} \tag{8}$$

where $h > 0$ is the fixed step length. When classified correctly, the motion direction is $-\dfrac{\mathbf{d}_k}{\|\mathbf{d}_k\|}$, which points toward the hole center. When misclassified, the direction is given by the unit vector $\mathbf{u}_s$, which satisfies $\|\mathbf{u}_s\| = 1$, but does not point toward the hole center.

The classification model satisfies the following assumptions:

1. The classification results at each step are independent and only depend on the current state.
2. There exists a constant $p > 0.5$ such that for all $d_k \neq 0$, the probability of correct classification satisfies
$$\mathbb{P}(\text{correct classification} \mid d_k) = p > 0.5.$$
3. Each of the four classes corresponds to a unique control direction, and the correct classification always corresponds to a direction that reduces $\|\mathbf{d}_k\|$.
4. During each assembly, the step size h can be chosen arbitrarily but must be a fixed value. As long as the search trajectory can cover the assembly clearance area around the hole, it can be considered as completing the task.

3.6.2. Convergence analysis in one-dimensional form

For the convenience of mathematical proof, we first consider the one-dimensional error case, where the error is given by

$$d_k \in \{0, h, 2h, 3h, \ldots\}.$$

When $d_k > 0$, a correct classification reduces the error to $d_k - h$, and a misclassification increases the error to $d_k + h$. Therefore, the system satisfies:

$$d_{k+1} = \begin{cases} d_k - h, & \text{with probability } p, \\ d_k + h, & \text{with probability } 1-p. \end{cases} \tag{9}$$

The state $d = 0$ is an absorbing state, indicating that the alignment process has been successfully completed.

Define

$$f(i) = \mathbb{P}(\text{starting from } d = ih, \text{ the process eventually reaches } 0).$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 22:56:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0263574726103695

It is clear that the boundary condition is

$$f(0) = 1.$$

According to the Markov property, $f(i)$ satisfies the recurrence relation:

$$f(i) = pf(i-1) + (1-p)f(i+1), \quad i \geq 1. \tag{10}$$

The general solution of Eq. (10) is

$$f(i) = A + B\left(\frac{p}{1-p}\right)^i. \tag{11}$$

When $p > 0.5$, we have

$$\frac{p}{1-p} > 1,$$

and if $B \neq 0$, $f(i)$ will diverge as $i \to \infty$, violating the probability constraint $0 \leq f(i) \leq 1$. Therefore, it must be that $B = 0$. From $f(0) = A = 1$, we obtain

$$f(i) \equiv 1, \quad \forall i \geq 0. \tag{12}$$

Equation (12) shows that when the correct classification probability $p > 0.5$, the one-dimensional system will reach the absorbing state $d = 0$. Even though misclassifications may occur, the overall bias toward the correct direction ensures that the random walk will not drift away indefinitely and will eventually move toward the hole center.

3.6.3. Convergence of two dimensional four classification system

The actual system is a two-dimensional error, which can be represented as:

$$d_k = (x_k, y_k).$$

The four classes correspond to the fixed directions of "front," "back," "left," and "right," with each correct class corresponding to a direction that reduces $\|\mathbf{d}_k\|$. In the actual assembly process, the error along one dimension is first reduced to zero, and then the error along the other dimension is adjusted. During single-dimensional adjustment, the classification results can be simplified into two classes: correct and incorrect. The position adjustment corresponding to the correct class reduces the error in that dimension, whereas the adjustment corresponding to the incorrect class may increase or maintain the error in that dimension.

On this basis, both dimensions follow a one-dimensional biased random walk model with a "probability of moving in the correct direction >0.5," and thus satisfy, respectively:

$$\mathbb{P}(|x_k| \to 0) = 1, \quad \mathbb{P}(|y_k| \to 0) = 1.$$

Since both events occur with probability 1, their intersection still has probability 1, thus:

$$\mathbb{P}(d_k \to 0) = 1.$$

That is, the two-dimensional system must inevitably reach the hole center region.

Even if there are recognition errors, as long as the overall bias in a single dimension leans toward the correct direction, the peg can still theoretically reach the center of the hole.

4. Experiments

4.1. Experiment environment

In order to verify the effectiveness of the proposed hole search strategy, a robot assembly experimental platform was built as shown in Figure 12. The platform consists of a six-degree-of-freedom collaborative robot (DOBOT CR5), a six-axis force/torque (F/T) sensor (KUNWEI KWR75B), a depth camera (Intel

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 22:56:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0263574726103695

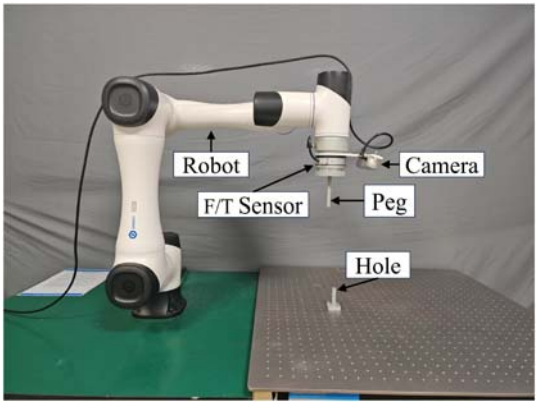


Figure 12. *Assembly experimental platform.*

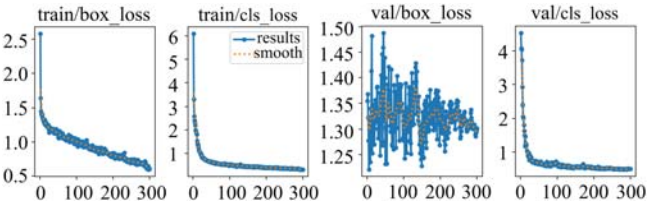


Figure 13. *Training loss curve of the YOLOv8n model.*

RealSense D455), a peg with a diameter of 8.97 mm, and a hole with a diameter of 9.1 mm. The assembly parts are fabricated from PLA material using 3D printing. For the peg-hole contact direction, admittance control is applied, with the desired contact force set to 10 N. The adjustment actions were implemented in two modes: position control mode and force control mode. In the position control mode, the adjustment step size for the peg-in-hole assembly experiment was set to 0.3 mm, and the step size for the USB-A connector insertion experiment was set to 0.5 mm. In force control mode, the robot relied on external force to drive motion, and the driving force for the peg-in-hole assembly experiment was set to 15 N, and the driving force for the USB-A connector insertion experiment was set to 18 N. In the experiment, the hole search time is defined as the time elapsed from the initial peg-hole contact to when the peg's insertion depth exceeds 1 mm.

4.2. Experimental results

To validate the effectiveness of the contact state classification model under fixed single-view conditions, the YOLOv8n model was employed for training and testing on the dataset. A dataset of 360 images was constructed using the aforementioned data collection method and classification rules, with a 17:3 ratio between the training and validation sets. The number of training epochs was set to 300, while all other parameters were kept at their default values. Training was automatically terminated when model performance no longer improved or the specified number of epochs was reached. The YOLOv8n training loss curves are shown in Figure 13. After training began, both train/box_loss and train/cls_loss curves dropped rapidly and gradually stabilized. The decline in the loss functions indicates that the prediction errors continuously decrease and the model performance steadily improves, suggesting that the training process is stable and the model converges well. On the validation set, the val/box_loss curve showed significant fluctuations. This can mainly be attributed to two factors. First, the target region accounts for only a small proportion of the image. Second, the boundary of the target region is not sufficiently clear, so approximate regions were used in the annotation process, which made it difficult for the model to learn precise boundary information. The rapid decrease in val/cls_loss curve indicates that the model can quickly learn effective classification features and has good generalization ability on the validation set.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 22:56:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0263574726103695

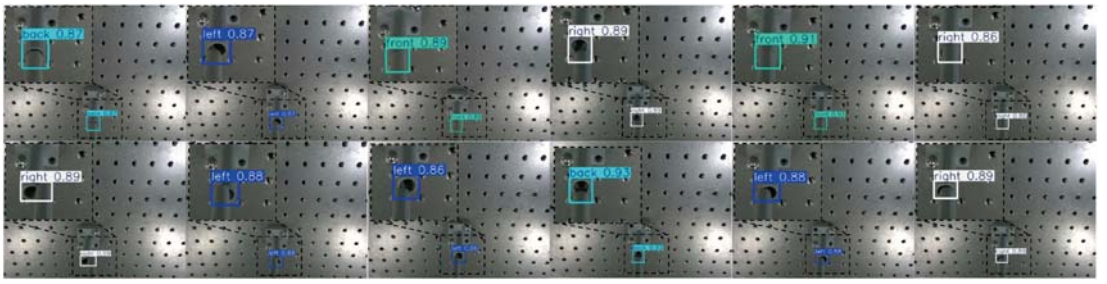


Figure 14. *Object detection results.*

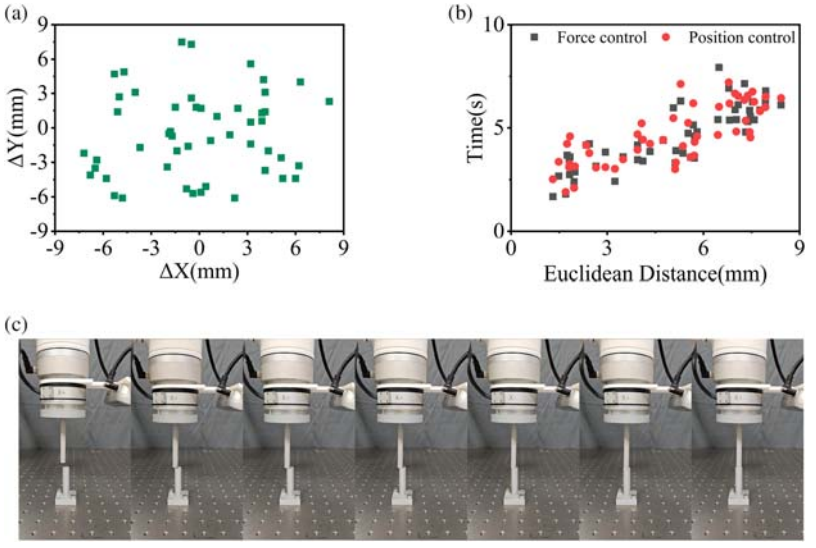


Figure 15. *Circular peg-in-hole assembly search experiment. (a) Random point distribution. (b) Relationship between Euclidean distance and hole search time. (c) Search process.*

The above results demonstrate that the trained model can accurately identify the contact state in images, indicating good classification performance. Although there are still some deviations in the localization of bounding boxes, this study primarily focuses on contact state recognition rather than precise localization; therefore, the current model is sufficient to meet the requirements of practical applications.

Figure 14 shows the results of contact state detection using the YOLOv8n model. Following data annotation and model training based on the classification criteria proposed in this paper, the model is able to accurately recognize four classes of contact states. Although the val/box_loss curve oscillates on the validation set, indicating that the detection box localization convergence is unstable, the model can still accurately locate the target region in actual use. The training curves and the test results together indicate that the model achieves good performance, which further verifies the feasibility of using image-based contact state classification to guide the search task.

While maintaining contact between the peg and the hole, 50 random initial contact positions were generated around the hole center, and their distribution is shown in Figure 15(a). This paper mainly investigates the hole search strategy in the assembly process and therefore focuses on the time consumption during the search phase [31]. For visualization, the random initial contact positions were characterized by the corresponding Euclidean distances between the peg and the hole, and the relationship between Euclidean distance and hole search time is shown in Figure 15(b). The results show that the hole search time generally increases with Euclidean distance, although noticeable fluctuations are observed. When the Euclidean distances are similar, the search time corresponding to some initial positions remains

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 22:56:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0263574726103695

Table I. *Comparative experiment of different search trajectories in the thin-walled hole search phase.*

Trajectory	Condition	Result
Spiral trajectory	$L \leq (R + R_{\text{Hole}})/2$	Success
	$R + r_{peg} > L > (R + R_{\text{Hole}})/2$	Loss of contact leads to search failure
Random trajectory	—	Loss of contact leads to search failure
Lissajous trajectory	—	Loss of contact leads to search failure
Ours	—	Success

significantly higher than that of neighboring random points. This is because the assembly clearance is small and the visual accuracy is limited, such that the peg sometimes requires more position adjustment steps before it can successfully enter the hole, thereby resulting in increased search time. Under the parameter settings used in this study, both the force control mode and the position control mode achieved a 100% hole search success rate. The average time consumption during the search phase is 4.56 s for the force control mode and 4.68 s for the position control mode. Figure 15 shows the flowchart of the peg-in-hole assembly search phase. After contact between the peg and the hole is established, the peg position is first adjusted in the left–right direction and then in the front–back direction.

When the left or right class is close to the boundary of the back class, the left or right class may be misclassified as the back class. This indicates that the effective region of the back class is larger than the predefined region. When the relative position between the peg and hole falls within this region, it will be classified as the back class, and the corresponding adjustment actions will be performed. However, the adjustment direction for these three contact states is always toward reducing the distance between the peg and the hole.

4.3. Comparative experiments

Blind search methods include various search trajectories, among which the spiral trajectory is the most widely used trajectory for thin-walled hole search tasks. In contrast, random trajectories, Lissajous trajectories, and other trajectories are more likely to cause loss of contact between the peg and the hole, thus leading to the failure of the search task. As shown in Table I, the spiral trajectory can only be applied to the thin-walled hole search phase under specific conditions, whereas the proposed search strategy is not constrained by these conditions. Next, the performance of the spiral search and the proposed search strategy will be compared experimentally under robot position control mode.

In the spiral search experiments, the initial contact distance between the peg and the hole was increased in steps of 0.5 mm until the conditions for using the spiral search were no longer satisfied. The parameters for the spiral trajectory are an initial radius of 0.1 mm, a maximum radius of 5 mm, a pitch of 0.08 mm, and an angular increment of 5° at each step. The search time of the spiral search was compared with that obtained from the above 50 randomly generated initial contact positions under position control mode. As shown in Figure 16(a), under the same Euclidean distance condition, the spiral search takes significantly longer than the proposed strategy and is more sensitive to initial distance errors, indicating that its performance is greatly affected by the initial position. Figure 16(b) shows the comparison of search trajectories between spiral search and the proposed strategy. As the initial distance increases, the traversal trajectory of spiral search significantly increases. Under the same initial position conditions, the search trajectory of the proposed strategy is shorter, so the search time is also shorter.

4.4. Rectangular hole search experiment

Due to the differences in visual features across different types of peg-in-hole assemblies, when this method is applied to new assembly objects, it typically requires the reacquisition of image data and the training of a corresponding model. To verify the applicability of the proposed strategy to rectangular peg-in-hole assembly, a USB-A connector was used as the experimental object in this study. Based

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 22:56:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0263574726103695

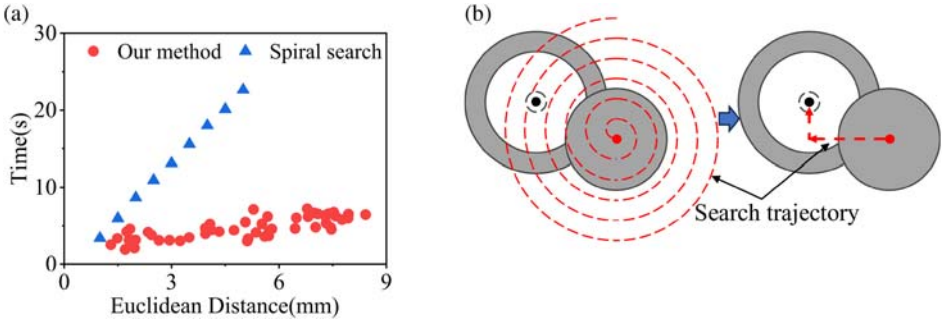


Figure 16. *(a) Comparison of hole search time. (b) Comparison of search trajectory.*

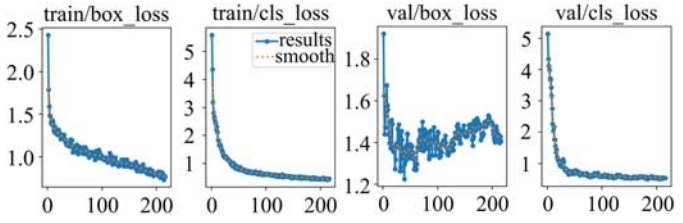


Figure 17. *Training loss curve of the YOLOv8n model on the USB-A connector contact state dataset.*

on the dimensions of the USB-A connector plug and socket, the data acquisition area was defined by establishing a coordinate system with the socket center as the origin and applying offsets of ±10 mm along the long axis and ±4 mm along the short axis. The collected data were annotated and used for training according to the defined contact-state classification rules. A total of 187 images were collected and divided into training and validation sets at a ratio of 9:1. The training parameters were the same as those used for the circular peg-in-hole task. The training loss and evaluation curves are shown in Figure 17. These curves demonstrate that the YOLOv8n model has successfully learned the relevant visual features.

While maintaining contact between the peg and the hole, 50 random initial contact positions were generated around the hole center, and their distribution is shown in Figure 18(a). To illustrate the relationship between search time and the initial contact position, the 50 random initial contact positions were represented by the Euclidean distances between the peg and the hole. Figure 18(b) shows the relationship between Euclidean distance and hole search time under the two control modes. The results indicate that the hole search time increases with increasing Euclidean distance, indicating that reducing the initial contact position error is beneficial for decreasing the time consumption during the search phase. Some random positions exhibit significantly longer hole search times than nearby positions, mainly because the assembly clearance is small and visual accuracy is limited, making it difficult to directly guide the peg into the assembly clearance region. Therefore, more position adjustment steps are required to complete the hole search process, and the additional adjustment steps increase the hole search time. The hole search success rate was 100% under both control modes. The average time consumption during the search phase is 3.98 s for the force control mode and 4.12 s for the position control mode. Figure 18(c) shows the search process in USB-A connector insertion assembly.

In rectangular peg-in-hole assembly, significant visual occlusion occurs. The visual detection on the left and right sides is severely affected, whereas the front and rear directions are largely unaffected. To compensate for the information loss caused by visual occlusion, a V-shaped trajectory is designed in the search phase of USB-A connector insertion assembly. When the contact state is identified as the front or back, lateral displacement is superimposed while the peg performs forward or backward position adjustment. When the contact state is identified as the left or right, the peg is adjusted only along the corresponding direction. By enlarging the area swept during the search process, the peg can reach regions that cannot be covered by a single adjustment direction, thereby eliminating the influence of visual occlusion on the search phase.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 22:56:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0263574726103695

Table II. *Comparison experiment of assembly success rate.*

Method	USB-A (%)	Circular peg-in-hole (%)
Hao et al. [32]	70	100
Gao et al. [33]	55	73
Li et al. [34]	84	100
Ours	100	100

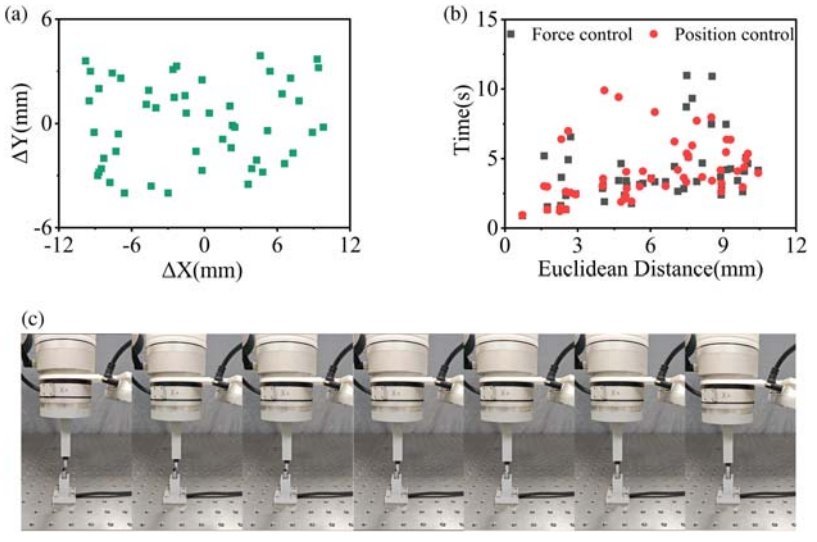


Figure 18. *(a) Random point distribution. (b) Relationship between Euclidean distance and hole search time. (c) Search process.*

4.5. Discussion

In practical applications, a small orientation deviation typically exists between the peg and the hole. Based on the dimensions of the circular peg-in-hole assembly, the critical tilt angle at which three-point contact occurs can be calculated. To verify the applicability of the proposed search strategy across the entire range, a boundary-testing approach is adopted. Within this range, the values -3° and 3° are selected as the tilt amounts about the two coordinate axes, and each experiment is repeated twice for every test condition. Since the method performs well in extreme cases, it can also be inferred that it has good robustness and adaptability at other angles in this range. The detection results of the contact state between the peg and the hole mainly depend on the distribution characteristics of the clearance under the camera's perspective. Even if the perspective changes slightly due to camera loading and unloading, the detection results remain stable, indicating that this method has good robustness.

According to the comparison results in Table II between the assembly method based on the proposed search strategy and other assembly methods, the assembly method using this strategy achieves a higher assembly success rate, indicating the good stability of the proposed strategy. The proposed search strategy ensures that the search process converges to the hole center, thereby significantly improving the overall reliability and success rate of the assembly.

5. Conclusion

This paper proposes a vision-guided hole search strategy. First, a contact-state classification model under fixed single-view conditions is constructed to support contact state detection and peg position adjustment. Then, the impact of camera mounting angle deviation and visual occlusion on the search

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 22:56:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0263574726103695

process is addressed by optimizing the search trajectory. Finally, a biased random walk model is used to demonstrate that a classification accuracy greater than 50% is sufficient to achieve successful hole search under the proposed strategy. The experimental results demonstrate that the proposed method is not only applicable to circular peg-in-hole assembly tasks but also exhibits good adaptability and stability in rectangular peg-in-hole assembly tasks. A 100% search success rate was achieved in both types of peg-in-hole assembly experiments. Compared with the spiral search, the proposed hole search strategy is less sensitive to initial position errors and achieves higher search efficiency. Future work will focus on integrating visual pose estimation techniques with the proposed hole search strategy to further enhance its autonomy and generalization capability in complex manufacturing environments.

Author contributions. Qinghua Lu and Weilin Chen conceived and designed the study. Xianda Wang and Shijie Tang performed data collection and experiments. Qinghua Zhang and Lufeng Luo performed the statistical analyses. Xianda Wang and Mingyou Chen wrote the manuscript. All authors read and approved the final manuscript.

Financial support. This work was supported by the National Natural Science Foundation of China (Grant Nos. 62273097, 52275010), Guangdong Basic and Applied Basic Research Foundation (Grant Nos. 2023A1515140090, 2021B1515120017, 2024A1515140133), Research Projects of the Universities in Guangdong Province (Grant No. 2023ZDZX3030), Key Technologies R&D Program of NEEII (Foshan) (Grant No. JBGS2025008), and Guangdong Provincial Key Laboratory of Industrial Intelligent Inspection Technology (Grant No. 2024B1212020011).

Competing interest. The authors declare no conflicts of interest exist.

Ethical approval. None.

References

[1] X. Song, P. Xu, W. Xu, B. Li and L. Qin, "Robot skill acquisition for precision assembly of flexible flat cable with force control," *Robotica* **42**(9), 2908–2923 (2024).
[2] Z. Wang, X. Li, H. Zhao, L. Shao, X. Ma, H. Zou and H. Ding, "Geometry and force guided robotic assembly with large initial deviations for electrical connectors," *IEEE Trans. Autom. Sci. Eng.* **22**, 8095–8107 (2024).
[3] Z. Wang, P. Li, H. Zhang, Q. Zhang, C. Ye, W. Han and W. Tian, "A binocular vision method for precise hole recognition in satellite assembly systems," *Measurement* **221**, 113455 (2023).
[4] E. Y. Puang, K. P. Tee and W. Jing, "Kovis: Keypoint-based Visual Servoing with Zero-Shot Sim-to-Real Transfer for Robotics Manipulation," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2020) pp. 7527–7533.
[5] M. Sileo, N. Capece, M. Gruosso, M. Nigro, D. D. Bloisi, F. Pierri and U. Erra, "Vision-enhanced Peg-in-Hole for automotive body parts using semantic image segmentation and object detection," *Eng. Appl. Artif. Intell.* **128**, 107486 (2024).
[6] J. C. Triyonoputro, W. Wan and K. Harada, "Quickly Inserting Pegs into Uncertain Holes Using Multi-View Images and Deep Network Trained on Synthetic Data," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2019) pp. 5792–5799.
[7] D. Wang, L. Ma, D. Huang, J. Lin and W. Lu, "Image-based adaptive stiffness visual servoing control for manipulators in feature space with prescribed performance," *IEEE/ASME Trans. Mechatron.* **30**(6), 7728–7739 (2025).
[8] D. Wang, J. Lin, L. Ma, D. Huang and Y. Wu, "Image-based visual-admittance control with prescribed performance of manipulators in feature space," *IEEE Trans. Ind. Electron.* **72**(5), 5060–5070 (2024).
[9] W. Wu, H. Zhou, Y. Guo, Y. Wu and J. Guo, "Peg-in-hole assembly in live-line maintenance based on generative mapping and searching network," *Robot. Auton. Syst.* **143**, 103797 (2021).
[10] H. Park, J. Park, D.-H. Lee, J.-H. Park, M.-H. Baeg and J.-H. Bae, "Compliance-based robotic peg-in-hole assembly strategy without force feedback," *IEEE Trans. Ind. Electron.* **64**(8), 6299–6309 (2017).
[11] H. Park, J. Park, D.-H. Lee, J.-H. Park and J.-H. Bae, "Compliant peg-in-hole assembly using partial spiral force trajectory with tilted peg posture," *IEEE Robot. Autom. Lett.* **5**(3), 4447–4454 (2020).
[12] H. Kang, Y. Zang, X. Wang and Y. Chen, "Uncertainty-driven spiral trajectory for robotic peg-in-hole assembly," *IEEE Robot. Autom. Lett.* **7**(3), 6661–6668 (2022).
[13] S. Wang, G. Chen, H. Xu and Z. Wang, "A robotic peg-in-hole assembly strategy based on variable compliance center," *IEEE Access* **7**, 167534–167546 (2019).
[14] M. Jokesch, J. Suchý, A. Winkler, A. Fross and U. Thomas, "Generic Algorithm for Peg-in-Hole Assembly Tasks for Pin Alignments with Impedance Controlled Robots," *Robot 2015: Second Iberian Robotics Conference: Advances in Robotics*, vol. **2** (Springer International Publishing, Cham, 2015) pp.105–117.
[15] Y. Mei, R. Li, Y. Tan, D. Zhu, T. Li and Z. Zhou, "Robotic assembly strategy with wrist force sense for narrow clearance peg-in-hole," *IEEE Trans. Autom. Sci. Eng.* **22**, 8709–8724 (2024).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 22:56:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0263574726103695

[16] M. Nigro, M. Sileo, F. Pierri, K. Genovese, D. D. Bloisi and F. Caccavale, "Peg-in-Hole Using 3D Workpiece Reconstruction and CNN-based Hole Detection," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2020) pp. 4235–4240.
[17] B. Ti, Y. Gao, M. Shi and J. Zhao, "Generalization of orientation trajectories and force–torque profiles for learning human assembly skill," *Robot. Comput.-Integr. Manuf.* **76**, 102325 (2022).
[18] Y. Liu, Z. Chen, H. Qiao and S. Gan, "Compliant peg-in-hole assembly for nonconvex axisymmetric components based on attractive region in environment," *Robotica* **41**(8), 2314–2336 (2023).
[19] Y.-L. Kim, B.-S. Kim and J.-B. Song, "Hole Detection Algorithm for Square Peg-in-Hole Using Force-based Shape Recognition," *IEEE International Conference on Automation Science and Engineering (CASE)* (IEEE, 2012) pp. 1074–1079.
[20] W. Wu, K. Liu and T. Wang, "Robot assembly theory and simulation of circular-rectangular compound peg-in-hole," *Robotica* **40**(9), 3306–3339 (2022).
[21] T. Tang, H. C. Lin, Y. Zhao, W. Chen and M. Tomizuka, "Autonomous Alignment of Peg and Hole by Force/Torque Measurement for Robotic Assembly," *IEEE International Conference on Automation Science and Engineering (CASE)* (IEEE, 2016) pp. 162–167.
[22] R. Xiao, C. Yang, Y. Jiang and H. Zhang, "One-shot sim-to-real transfer policy for robotic assembly via reinforcement learning with visual demonstration," *Robotica* **42**(4), 1074–1093 (2024).
[23] S. Li, B. Li and J. Yu, "Integrating admittance control and residual reinforcement learning for robotic assembly with visual and force feedback," *Robotica* **44**(1), 37–51 (2025).
[24] N. Zhang, Y. Zhao, M. Yang and S. Dai, "Image segmentation-driven sim-to-real deep reinforcement learning framework for accurate peg-in-hole assembly," *Robotica* **43**(8), 2783–2802 (2025).
[25] J. Song, Q. Chen and Z. Li, "A peg-in-hole robot assembly system based on Gauss mixture model," *Robot. Cim-INT Manuf.* **67**, 101996 (2021).
[26] H. Chu, T. Zhang, Y. Zou and H. Sun, "Assemble like human: A multi-level imitation model learning human perception-decision-operation skills for robot automatic assembly tasks," *Robot. Cim-INT Manuf.* **93**, 102907 (2025).
[27] Z. Chen, S. Xie and X. Zhang, "Position/force visual-sensing-based robotic sheet-like peg-in-hole assembly," *IEEE Trans. Instrum. Meas.* **71**, 1–11 (2021).
[28] E. Mobedi, N. Villa, W. Kim and A. Ajoudani, "An Adaptive Control Approach to Robotic Assembly with Uncertainties in Vision and Dynamics," *29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)* (IEEE, 2020) pp. 144–150.
[29] B. Liu, Y. Li, X. Feng and P. Lian, "BEW-YOLOv8: A deep learning model for multi-scene and multi-scale flood depth estimation," *J. Hydrol.* **645**, 132139 (2024).
[30] L. Nie, B. Li, F. Jiao, W. Lu, X. Shi, X. Song, Z. Shi, T. Yang, Y. Du and Z. Liu, "EVIT-YOLOv8: Construction and research on African Swine Fever facial expression recognition," *Comput. Electron. Agric.* **227**, 109575 (2024).
[31] J. Jiang, L. Yao, Z. Huang, G. Yu, L. Wang and Z. Bi, "The state of the art of search strategies in robotic assembly," *J. Ind. Inf. Integr.* **26**, 100259 (2022).
[32] P. Hao, T. Lu, S. Cui, J. Wei, Y. Cai and S. Wang, "Meta-residual policy learning: Zero-trial robot skill adaptation via knowledge fusion," *IEEE Robot. Autom. Lett.* **7**(2), 3656–3663 (2022).
[33] W. Gao and R. Tedrake, "kpam 2.0: Feedback control for category-level robotic manipulation," *IEEE Robot. Autom. Lett.* **6**(2), 2962–2969 (2021).
[34] L. Jin, Y. Men, R. Song, F. Li, Y. Li and X. Tian, "Robot skill generalization: Feature-selected adaptation transfer for peg-in-hole assembly," *IEEE Trans. Ind. Electron.* **71**(3), 2748–2757 (2023).

Cite this article: X. Wang, Q. Lu, W. Chen, S. Tang, Q. Zhang, L. Luo and M. Chen (2026). "Hole search strategy for robotic peg-in-hole assembly under initial positioning uncertainty", Robotica **44**, 1703–1720. https://doi.org/10.1017/S0263574726103695

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 22:56:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0263574726103695