IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

Multi-strategy Enhanced Particle Swarm Optimization for Variable Curvature Path Planning in Flexible Needle Insertion

Yanding Qin, *Member, IEEE*, Jianing Teng, Chao Wen, Ge Fang, *Member, IEEE*, Hongpeng Wang, *Member, IEEE*, and Jianda Han, *Member, IEEE*

***Abstract*— Flexible needles provide enhanced adaptability for navigating puncture pathways and avoiding obstacles when compared to conventional rigid needles. However, developing a three dimensional (3D) curved path for flexible needle is challenging, particularly in achieving both effective obstacle avoidance and precise targeting. To this end, we proposed an improved particle swarm optimization-based path planning approach by incorporating good point set initialization and heuristic multi-mutation strategy. Such incorporation greatly enhanced the algorithm's global search capability while ensuring fast convergence speed. In addition, 3D biarc curve fitting was employed to develop a kinematically reachable path for bevel tip needles. Obstacle-avoidance simulations conducted demonstrate the superior performance of proposed method against state-of-the-art algorithms in the aspect of path length and distance to obstacles, repeatability and local minima trap avoidance. Needle puncturing experiments performed using duty cycling control achieved a small curvature radius of 49.6 mm and targeting errors of less than 4 mm. This algorithm facilitates efficient variable curvature path planning for flexible needles, ensuring precise targeting while effectively avoiding obstacles.**

***Index Terms*—Flexible needle path planning, biarc curve, particle swarm optimization, good point set, mutation solution.**

I. INTRODUCTION

In minimally invasive surgery, needle puncture is commonly used for various medical tasks, including biopsy sampling, drug injection, tissue ablation, and brachytherapy. This technique offers the advantages of minimal trauma and enables rapid patient recovery, making it a preferred option for clinical diagnosis and treatment. However, conventional rigid needles used in clinical practice can only achieve a straight trajectory from the surgical keyhole to targets [1]. The lack of active bending capability makes it difficult to avoid critical structures along the puncture pathway.

To address the limitations of rigid needles in puncture, a bevel-tipped flexible needle design using nickel-titanium alloy was proposed [2]. Through appropriate rotation and insertion techniques, this flexible needle can navigate around obstacles to effectively reach the desired target [3]. While flexible needles provide increased agility, they also introduce substantial complexities in path planning and control during puncture interventions. When the puncture process of flexible needles relies heavily on the physician's experience, there is an increased risk of severe complications for patients, including secondary bleeding. To enhance the accuracy and safety of flexible needle punctures, it is crucial to pre-plan a reference path for guidance to assist surgeons.

Recently, flexible needle path planning has attracted significant attention from researchers. The current approaches can be divided into two main categories: numerical methods and exploration-based methods. Numerical methods usually involve defining an objective function, which is then optimized using numerical computation to determine the optimal path. Duindam et al. employed inverse kinematics to develop both two-dimensional (2D) and three-dimensional (3D) paths for obstacle avoidance, fixing the target point and the joints of the puncture trajectory [4]. Lee and Park applied the geometric relationships of circular arcs to plan a single entry point that can reach multiple targets [5]. Wang et al. proposed a spring-damping model that represents the 3D puncture trajectory as a geometric combination of multiple 2D planes. They optimized the number of needle rotations and the length during the puncture process to reduce tissue damage [6]. However, solving the complex optimization problem demands significant computation time, particularly in a 3D environment. Huo et al. proposed a simplified kinematic model of the flexible needle and used particle swarm optimization (PSO) to optimize the sequence of puncture lengths and rotation angles [7]. Cai et al. utilized PSO to develop a path planning method, and analyzed the impact of parameters, such as the needle tip angle, on the radius of the circular arcs during the path planning process [8]. However, the trajectory of the flexible needle is still modeled with a fixed curvature, which considerably limits the flexibility of the needle's movement. Besides, standard PSO algorithm can converge at local minima [9].

Alternatively, several exploration-based algorithms for flexible needle path planning have been proposed. Patil and Alterovitz introduced the rapidly-exploring random tree (RRT)

Manuscript received: February 19, 2025; Revised: April 8, 2025; Accepted: June 10, 2025. This paper was recommended for publication by Editor Pietro Valdastri upon evaluation of the Associate Editor and Reviewers' comments.

This work was supported in part by the National Key Research and Development Program of China under Grant 2022YFB4702800, in part by the Shenzhen Science and Technology Program under Grant KQTD202108110 90143060, and in part by the Sustainable Development Science and Technology Special Project of Shenzhen under Grant KCXFZ202307311 00900002. (*Corresponding authors: Ge Fang; Jianda Han*)

All the authors are with the College of Artificial Intelligence, Nankai University, Tianjin 300350, China, and also with the Institute of Intelligence Technology and Robotic Systems, Shenzhen Research Institute of Nankai University, Shenzhen 518083, China (e-mail: qinyd@nankai.edu.cn; fangge@nankai.edu.cn; hanjianda@nankai.edu.cn).

Digital Object Identifier (DOI): see top of this page.

algorithm, which utilizes variable curvature arc trajectories coupled with a reachability-guided sampling heuristic method. This approach significantly enhances the planning speed [10]. Bernardes et al. developed an online motion planning method that employs duty-cycled rotation and RRT-based replanning for steerable needles in dynamic 3D environments. This method effectively addresses uncertainties such as tissue deformation and patient movement [11]. Aghdam and Liu proposed a multi-target path planning approach during the RRT extension process. Their method takes advantage of the flexible needle's capability to retract along its original trajectory during actual insertion, which aims to improve trajectory flexibility in static environments [12]. Zhang et al. presented an enhanced RRT* algorithm for 3D path planning of robot-assisted flexible needle insertion into multilayer tissues. The method enhances path smoothness and safety by adjusting parameters for different tissue types during insertion and by incorporating artificial potential fields to navigate around obstacles [13]. Huang and Zhang proposed a variable curvature path planning approach for robot-assisted flexible needle insertion based on an improved Bi-RRT algorithm. The method uses target-oriented and bidirectional-search strategies to boost search efficiency and minimize redundant nodes, thus improving trajectory control accuracy [14]. These algorithms enhance the branching tree structure with arc-based stitching to accommodate the flexible needle trajectory, thereby continuously improving the efficiency of the algorithm. However, the intrinsic randomness associated with RRT algorithms may result in local minima. Furthermore, RRT algorithms may generate paths with short arcs, which complicates the needle motion control.

This paper aims to propose a method that enables variable curvature path planning for flexible needles with improved global search capability. Hybrid mutation strategy is integrated during particle swarm evolution to escape from local minima. Besides, instead of randomly generating initial particles, good point set initialization is employed in our method to enhance the diversity of particle swarm distribution. In addition, 3D biarc path fitting is incorporated for path refinement to meet the kinematic constraint of flexible needle. The major contributions are summarized as follows:

1) A multi-strategy enhanced optimization approach based on PSO, incorporating good point set initialization and hybrid mutation. This approach facilitates variable curvature path planning with enhanced global search capabilities while ensuring a high level of computational efficiency.
2) A 3D biarc mathematical model for the motion trajectory of flexible needle to fit the pre-planned path, following the kinematic characteristics of the needle.
3) Simulation tests for obstacle avoidance in 3D space to validate and compare the performance of proposed method. Additionally, flexible needle puncture experiments, performed with duty cycling control, demonstrate precise target reaching along the planned path.

The remainder of this paper is organized as follows. Section II introduces the needle puncture method and the unicycle model. Section III presents the Improved-PSO (IPSO) algorithm with good point set and hybrid mutation strategies. This section also introduces a biarc mathematical model designed to fit the path, ensuring it aligns with the needle's kinematic requirements. Section IV offers a comparative analysis of algorithm simulations, details the needle puncture experiments, and discusses the results. Finally, section V concludes with a summary and discusses potential future directions for research.

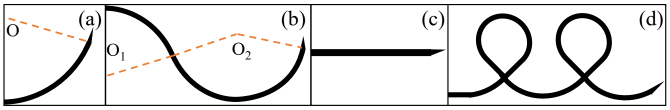

Fig. 1. Puncture trajectories of flexible needle: (a) C-shape with single arc; (b) S-shape with multiple arcs; (c) Straight line; (d) Spiral path.

II. PROBLEM STATEMENT

A. The Puncture Method

A flexible needle with an asymmetrical bevel tip experiences uneven lateral forces when penetrating soft tissues, resulting in deflection of the needle tip and bending of the shaft [15], [16]. It is generally recognized that the puncture mechanism of a flexible needle operates with two degrees of freedom (DOF): translation (feed) and rotation [17]. The method for puncture control, known as "duty-cycling", enables the flexible needle to generate trajectories with varying curvature radii by adjusting the insertion and rotation timing within a single cycle [18], [19]. During each cycle, the needle punctures continuously, with rotation occurring only for a portion of the time (τ_{rot}), while the remaining time is dedicated purely to insertion (τ_{stop}). Assuming constant insertion and rotation velocities, the duty cycle (α) is defined by

$$\alpha = \tau_{rot} / T = \tau_{rot} \Big/ \left(\tau_{rot} + \tau_{stop} \right) \tag{1}$$

As the duty-cycling control progresses, the curvature κ of the flexible needle can be defined by

$$\kappa = \kappa_{max} (1 - \alpha) \tag{2}$$

where κ_{max} represents the natural curvature of the flexible needle obtained using the stop-and-run puncture method, which is also its maximum curvature.

By employing the duty-cycling method, we can theoretically achieve circular arc trajectories with radii ranging from $1/\kappa_{max}$ to positive infinity. The flexible needle with a bevel tip can achieve trajectories in the form of multiple arcs, spirals, and a straight line, as illustrated in Fig. 1.

B. Needle Kinematics with Unicycle Model

To analyze the kinematic behavior of the needle tip, we make the following assumptions: the needle tip rotates in synchrony with the base [20]; no significant tissue deformation occurs during the puncture process; the needle shaft completely aligns with the needle tip and does not experience any radial displacement during movement [21].

Fig. 2 illustrates the unicycle model of a flexible needle as it punctures tissue, with W_{xyz} representing the world coordinate system. F_1 and F_2 denote the coordinate systems at the needle tip. Following the feed and rotation operations, the flexible needle moves from the origin of F_1 to the origin of F_2 controlled by the inputs v (velocity) and ω (angular velocity).

The actual curvature κ $(1/r)$ of the path traversed by the needle tip is influenced by tissue characteristics. The kinematics model can be described as follows:

$$V^b = \left[v^{\mathrm{T}} \ \omega^{\mathrm{T}} \right]^{\mathrm{T}} = v(t)^{\mathrm{T}} V_1 + \omega(t)^{\mathrm{T}} V_2 \tag{3}$$

$$V_1 = \left[e_3^{\mathrm{T}} \ \kappa e_1^{\mathrm{T}} \right]^{\mathrm{T}}, V_2 = \left[0^{\mathrm{T}} \ e_3^{\mathrm{T}} \right]^{\mathrm{T}} \tag{4}$$

where V^b represents the motion state of the needle tip in the coordinate system of the needle body, e_1, e_2, e_3 are unit vectors. The transformation matrix G_{wb} of the body coordinate system relative to the world coordinate system is shown as:

$$G_{\mathrm{wb}} = \begin{bmatrix} R_{\mathrm{wb}} & p_{\mathrm{wb}} \\ 0 & 1 \end{bmatrix} \in \mathrm{SE}(3) \tag{5}$$

where R_{wb} and p_{wb} respectively represent the rotation and translation of coordinate system F_1 relative to coordinate system F_2. Consequently, the movement of a flexible needle tip can be defined as follows:

$$G_{\mathrm{wb}}(t) = G_{\mathrm{wb}}(0) e^{\hat{V}^b t} \tag{6}$$

where ˆ denotes the wedge operator that transforms R^3 to the Lie algebra of SE(3), $G_{\mathrm{wb}}(0)$ is the initial configuration of the flexible needle, t represents the current time.

The unicycle model for bevel-tip needles implies that the needle path is formed by consecutive tangential-continuous arcs, which needs to be satisfied to ensure kinematical feasibility of the path.

III. ALGORITHM OPTIMIZATION STRATEGY

PSO describes the foraging behavior of bird flocks or fish schools in nature [22]. It is primarily illustrated through two key terms: $\boldsymbol{x}_{best}$ and $\boldsymbol{G}_{best}$. $\boldsymbol{x}_{best}$ represents the optimal solutions for individual particles, while $\boldsymbol{G}_{best}$ represents the optimal solution for the entire swarm of particles. The evolution equations for the velocity ($\boldsymbol{v}$) and position ($\boldsymbol{x}$) of the particle swarm over successive iterations, t, can be expressed as:

$$\begin{aligned} \boldsymbol{v}_k(t+1) &= w \cdot \boldsymbol{v}_k(t) + \mathrm{c}_1 \xi_1 \left(\boldsymbol{x}_{\mathrm{best},k}(t) - \boldsymbol{x}_k(t) \right) \\ &\quad + \mathrm{c}_2 \xi_2 \left(\boldsymbol{G}_{\mathrm{best}}(t) - \boldsymbol{x}_k(t) \right) \\ \boldsymbol{x}_k(t+1) &= \boldsymbol{x}_k(t) + \boldsymbol{v}_k(t+1) \\ \boldsymbol{x}_k &= \left(x_k^1, x_k^2, \cdots, x_k^j \right)^T \\ (k &= 1, 2, \ldots, m) \text{ and } (j = 1, 2, \ldots, n) \end{aligned} \tag{7}$$

where k denotes the indices of the particle within a swarm of size m, j represents the dimension index of the workspace. The constants c_1, c_2 are both set to 1.5. ξ_1 and ξ_2 are random numbers within the range of 0 to 1. w is the time-varying dynamic inertia weight given as follows:

$$w = 1 \Big/ \left(1 + e^{c_3 \cdot (t - 0.5\lambda)} \right) \tag{8}$$

where λ is the maximum number of iterations. For our path planning application, each particle is a path composed of N control points. The mechanism of jointly adjusting particles' velocities and positions based on both personal best and global best values in the particle swarm can lead to convergence towards local optima [9]. To enhance the algorithm's global search capability while effectively balancing exploration and exploitation, we propose the following strategies:

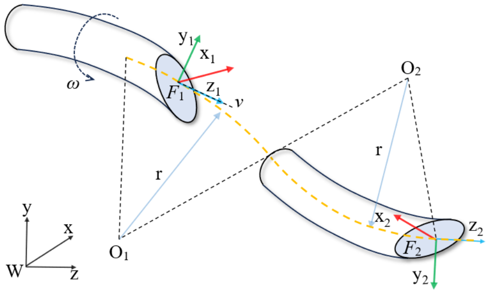

Fig. 2. The unicycle model of the bevel-tip flexible needle punctured in tissue.

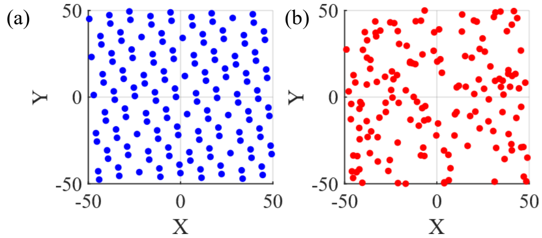

Fig. 3. The comparison of point set distributions initialized in 2D workspace with: (a) good point set, (b) random point set.

A. The Good Point Set (GPS)

Most swarm-based algorithms randomly generate initial particle positions $\boldsymbol{x}_{ori}$ and velocities $\boldsymbol{v}_{ori}$. However, the distribution pattern of initial particles can significantly affect the speed of global search and the quality of optimal solution [23]. The impact on the algorithm's performance is ultimately reflected in its iterative process. As a result, methods that leverage good point set theory for initializing population distributions have gained considerable attention in practical application. Let G_n be the unit cube of a n-dimensional Euclidean space, a point is denoted as $r_g = (r_{g1}, r_{g2}, \ldots, r_{gn}) \in G_n$. The GPS $S_m(k)$ can be defined as: $S_m(k) = \{(\{r_{g1}^{(m)} \cdot k\}, \{r_{g2}^{(m)} \cdot k\}, \ldots, \{r_{gn}^{(m)} \cdot k\})\}$, where $\{r_{gj}^{(m)} \cdot k\}$ is the decimal part of $r_{gj}^{(m)} \cdot k$, and m is the number of good points. The deviation of the GPS is calculated as: $\phi(m) = \mathrm{C}(r_g, \varepsilon) \cdot m^{-1+\varepsilon}$, where ε is an arbitrary positive number, and $\mathrm{C}(r_g, \varepsilon)$ is a constant associated with ε and r_g. Take $R_{gj} = \{2\cos(2\pi \cdot j/p), 1 \leqslant j \leqslant n\}$, where p is the smallest prime satisfying $p \geqslant 2n+3$, we can obtain r_{gj} as $mod(R_g, 1)$. Finally, r_g is mapped to the uniform point set x_k^j in the corresponding workspace, namely $x_k^j = (ub_j - lb_j) \cdot S_m(k) + lb_j$, where lb_j and ub_j respectively represent the lower and upper bounds of the workspace for dimension j.

Fig. 3 shows the good point set and random point set, both with a population size of 150 in a 2D space. The good point set $\boldsymbol{x}_{ori}$ achieves a more uniform distribution within the population space, thereby enhancing the population diversity. We utilize this method to improve the algorithm's global search capabilities while ensuring a fast convergence speed.

B. The Mutation Solutions (TMS)

During the iteration process, random mutation of the population's positions aids in avoiding being trapped in local minima. Inspired by the evolutionary mechanism of Whale Optimization Algorithm (WOA) [24], a hybrid mutation strategy is proposed. This strategy consists of interactive selection-based clustering center mutation and spiral function mutation, along with an independent Cauchy-Gaussian mutation perturbation method. These mutation behaviors are executed at each iteration. The detailed mutation steps are described as follows.

The Cauchy mutation enhances population diversity and accelerates the search for the optimal solution, while the Gaussian mutation strengthens the algorithm's local search capability, helping to prevent getting trapped in local minima. At each iteration of the algorithm, a fixed mutation rate of 10% (c_4) is applied. Within this probability range, the globally optimal individual, $\boldsymbol{G}_{best}$, in the population undergoes Gaussian and Cauchy mutations with equal probability of 50% (c_5), as represented in the following formula:

$$\boldsymbol{x}_{new} = \begin{cases} \boldsymbol{G}_{best} + 0.1 \cdot R_{Gsu}, 0.5 \leqslant \zeta < 1 \ \text{(a)} \\ \boldsymbol{G}_{best} + 0.1 \cdot R_{Cau}, 0 \leqslant \zeta < 0.5 \ \text{(b)} \end{cases} \tag{9}$$

where ζ represents probability. R_{Cau} is the Cauchy random number, R_{Gsu} is the Gaussian random number represented by:

$$R_{Cau} = \tan\left(\pi\left(\xi_3 - 0.5\right)\right), R_{Gsu} = 1/\sqrt{2\pi} \cdot e^{-0.5\xi_3^2} \tag{10}$$

where ξ_3 is a random number and $\xi_3 \in (0, 1)$.

Apart from the independent mutation method mentioned above, we applied a hybrid strategy that combines clustering center and spiral mutation for particles with poor fitness. This strategy is inspired by the encircling and bubble-net attacking behaviors exhibited by whale pods in the WOA.

Firstly, mutation based on clustering center is conducted. The population size m remains constant at 150. During each iteration, the particles are arranged in ascending order of fitness value and sequentially divided into three categories. The top 20 particles are classified as elite particles ($\boldsymbol{x}_{elite}$) to calculate the coordinates of the center particle. Particles ranked from 21 to 49 are designated as mediocre particles. Mutation operations are applied to the inferior particles, starting from rank 50 and onward. The new particle $\boldsymbol{x}_{new}$ is given as follows:

$$\boldsymbol{x}_{new} = \boldsymbol{x}_{center} + 2\xi_4 - 1 \tag{11}$$

where $\boldsymbol{x}_{center}$ refers to the center elite particles, derived by averaging all elite particles at the current iteration. $\xi_4 \in (0, 1)$ is a random number.

Secondly, spiral function mutation is employed. In this mutation method, a spiral function similar to the bubble-net at tacking behavior in the WOA is selected. The spiral function can be described as follows:

$$y = e^{(-\frac{3t}{\lambda}-1)\cdot\xi_5+1} \cdot \cos\left(2\pi\left((-\frac{3t}{\lambda}-1)\cdot\xi_5+1\right)\right)+1 \tag{12}$$

where t and λ represent the current iteration count and the total iteration counts respectively; ξ_5 is a random number and $\xi_5 \in (0, 1)$. Then a new particle $\boldsymbol{x}_{new}$ can be obtained as follows:

$$\boldsymbol{x}_{new} = y \cdot \boldsymbol{x}(k), 50 < k \leqslant m \tag{13}$$

where $\boldsymbol{x}(k)$ represents one of the inferior particles. The hybrid use of the clustering center mutation and spiral function mutation requires inferior particles to mutate according to the dynamic mutation rate, δ, as specified below:

$$\delta = 1 - t/\lambda \tag{14}$$

C. 3D Biarc Path Fitting

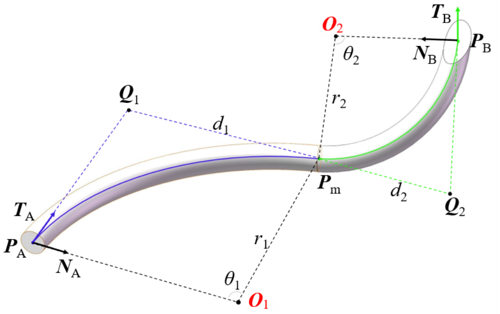

Fig. 4. Geometrical modelling of the flexible needle trajectory with two consecutive arcs in 3D.

The trajectory formed by the point set obtained through the above algorithm does not accurately match the actual motion of flexible needle, i.e. the unicycle model, which must meet the following requirements: The trajectory of a flexible needle is composed of multiple arc segments, with the radius of each arc being larger than the needle's natural curvature radius; The trajectory must satisfy G^1 continuity. To achieve this, a 3D biarc mathematical model for the flexible needle is developed to describe the relationship between two arcs, as illustrated in Fig. 4. The points $\boldsymbol{P}_A$ and $\boldsymbol{P}_B$ are generated from the IPSO algorithm. After the needle's movement, two circular arcs with radii r_1 and r_2 are smoothly connected at point $\boldsymbol{P}_m$. The tangential directions of $\boldsymbol{P}_A$ and $\boldsymbol{P}_B$ are defined as $\boldsymbol{T}_A$ and $\boldsymbol{T}_B$, with the normal vectors denoted as $\boldsymbol{N}_A$ and $\boldsymbol{N}_B$. $\boldsymbol{O}_1$ and $\boldsymbol{O}_2$ represent the center points of the two arcs, with cover angles of θ_1 and θ_2 respectively. By extending lines from point $\boldsymbol{P}_m$ at both tangential directions and intersecting them with the extended lines of $\boldsymbol{T}_A$ and $\boldsymbol{T}_B$, we can obtain points $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$. The lengths of Q_1P_m and Q_2P_m, d_1 and d_2, can be determined as follow:

$$d_2 = \rho d_1, d_1 = \left(-\gamma_1 \pm \sqrt{\gamma_1^2 - 4\gamma_2\gamma_0}\right) \Big/ 2\gamma_2 \tag{15}$$

where $\gamma_2=\rho(\boldsymbol{T}_A \cdot \boldsymbol{T}_B-1)$, $\gamma_1=(\boldsymbol{P}_A-\boldsymbol{P}_B)\cdot(\boldsymbol{T}_A-\rho\boldsymbol{T}_B)$, and $\gamma_0=(\boldsymbol{P}_B-\boldsymbol{P}_A)\cdot(\boldsymbol{P}_B-\boldsymbol{P}_A)/2$. Then the length of every arc segment, l, can be obtained by $l = \theta \cdot r$. Once the points set is obtained by the IPSO algorithm, the variables to be optimized are the set of tangent vectors, $\boldsymbol{T}$, and the set of ratios, ρ, which enables flexible generation of multiple circular arc parameters in steerable needle trajectories for experimental validation.

IV. SIMULATIONS AND EXPERIMENTS

A. Performance Comparison of the Algorithm, IPSO

To validate the performance of this multi-strategy improved algorithm, we conducted simulation studies using Matlab 2022a on a PC with Intel i9-12900HX 2.30GHz CPU and 24-GB RAM. Initially, the 3D simulation environments with single obstacle and multiple obstacles were constructed. The simulation workspace features a size of 80 mm × 70 mm × 50 mm with the origin set at (0, 0, 0) and the target point set to (50, 45, 25). The single obstacle is a sphere located at (25, 22.5, 12.5) with a radius of 8 mm. In our simulation tests, we implemented PSO, WOA, and the hybrid WOA-PSO algorithm (HWPSO) [25] for comparison. To ensure fairness in the comparison, we fixed the control parameters for these algorithms: the max iteration (λ), the number of population size (m) and the number of optimized control points (N). The objective function, i.e., the fitness f, was defined as follows:

$$f = L\left(\omega_0 + \omega_1 \cdot V\right) \tag{16}$$

where the coefficients are defined as: ω_0=1, ω_1=500. The variable L represents the total length of the path, which is composed of M discrete points interpolated from the N control points. The path length is calculated as the cumulative sum of the distances between adjacent points. V denotes the risk value associated with the distance between the path and obstacles.

To ensure safety, we introduced a threshold Ψ of 5 mm, which can be adjusted based on the specific environment. This threshold indicates that if the path remains at least 5 mm away from the surface of a spherical obstacle, the objective function f will not be penalized by V. This relationship can be expressed using the following formula.

$$V = \sum_{i=1}^{M}\sum_{j=1}^{6} \max\left(1-\left(\sqrt{\left(\boldsymbol{x}_i - \boldsymbol{x}_{obs,j}\right)^2} - \Psi\right)\Big/ r_{obs,j}, 0\right) \tag{17}$$

As in TABLE I, the influences of core parameters, i.e. the maximum iteration count λ, population size m, number of optimized control points N and the dynamic inertia weight constant c_3, were studied with control variable method to facilitate algorithm customization. We varied these parameters and conducted simulations in a multi-obstacle environment. The performances were evaluated in terms of processing time, final fitness value, and path smoothness. The path smoothness (μ) was quantified using the sum of squared forward second-order differences of the 3D coordinates.

Four parametric studies were conducted. In the first study, when varing N (5-9) with fixed parameters (λ=100, m=150, c_3=0.095), it was observed that the algorithm achieved the best results in terms of minimum iteration time, fitness value, and smoothness when N was set to 5, outperforming the settings for 7 and 9. In the second study, we increased c_3 from 0.055 to 0.135, with the parameters λ, m, and N set at 100, 120 and 5, respectively. When c_3 was set to 0.095, all three metrics outperform that with c_3 valued at 0.055 and 0.135. This indicates that a dynamic inertial weight constant of 0.095 provides a balanced trade-off between global exploration and local exploitation. In the third study, with m varied among 150, 325 and 500, the parameters λ, c_3, and N were fixed at 120, 0.045 and 5, respectively. It was found that increasing the population size substantially increased computational time with only minimal improvement in fitness value and path smoothness. This suggests that a population size of 150 is sufficient for the validated 3D environment. In the fourth study, λ was varied while m, c_3, and N remained fixed at 100, 0.045 and 5. It was observed that the number of iterations is positively correlated with the algorithm's runtime. With an iteration count of 50, the fitness value remained high at 87.98 and showed a decreasing trend, indicating that this iteration count is insufficient. When the iteration count was increased from 50 to 150, smoothness improved from 0.4 to 0.29. However, the enhancement in fitness value is minimal, with the computation time significantly increased. Therefore, an iteration count of 100 is a balanced choice. In summary, for IPSO, we recommend setting the values of N, c_3, m and λ to 5, 0.095, 150 and 100, respectively.

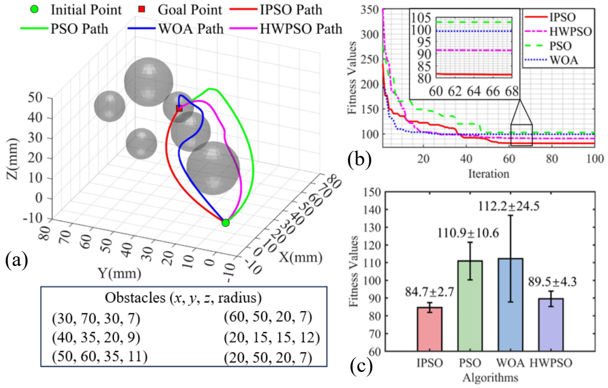

Fig. 5. Path planning simulation in multi-obstacle environment using various algorithms. (a) Paths generated. (b) Convergence curves. (c) Statistics from 100 repetitive experiments.

TABLE I. SIMULATION RESULTS WITH DIFFERENT IPSO PARAMETERS

Parameters	Designated value	Iteration time (s)	Fitness value	Smoothness μ
N	**5**	**11.69**	**83.95**	**0.31**
	7	12.14	84.92	0.53
	9	12.57	91.09	0.68
c_3	0.055	10.23	89.50	0.47
	0.095	**9.93**	**81.10**	**0.23**
	0.135	10.39	83.97	0.33
m	**150**	**15.39**	**83.45**	**0.35**
	325	23.48	83.50	0.34
	500	33.75	83.23	0.32
λ	50	5.08	87.98	0.40
	100	**9.48**	**81.36**	**0.37**
	150	14.42	81.20	0.29

To provide a thorough comparison of the proposed algorithm with prior arts, we conducted simulations involving both single and multiple obstacles for each algorithm. In Fig. 5, we presented the fitness variation of these algorithms over iterations in multi-obstacle setting, as well as the path visualization in the 3D environment. Through 100 repetitive trials for each algorithm, we summarized the final fitness values using a line chart, with statistic metrics such as maximum value, minimum value, and standard deviation (S_{td}) calculated. The data from the repetitive experiments are presented in TABLE II.

From Fig. 5 (b), it can be observed that the fitness value of IPSO rapidly declines and converges to an optimal value of 81.22 by the 60th iteration. At this point, the risk function's penalization is zero, meaning that the path length is equal to the fitness value, which is the lowest among all the compared algorithms. Additionally, the trajectory generated by IPSO is the smoothest compared to the other algorithms. While the HWPSO algorithm, which incorporates elements from the WOA algorithm, exhibits a slightly faster convergence speed, it still fails to reach the optimal solution under the same conditions, similar to the other algorithms. IPSO demonstrated superior performances, particularly in multi-obstacle scenario. In summary, we can conclude the following points:

1) The IPSO algorithm enables effective exploration of optimal solutions. IPSO achieves the lowest fitness value of 84.66 across multiple experiments, which is 11.5%, 17.4% and 1.3% less than that of WOA (91.55), PSO (98.11), and HWPSO (82.09), respectively.
2) IPSO offers excellent capacity for obstacle avoidance and

is less prone to getting trapped in local minima. Its maximum fitness value is 90.69, which is lower by 59.3%, 50.3% and 10.5% compared to the WOA (222.73), PSO (182.31), and HWPSO (101.38) algorithms, respectively.

3) The IPSO algorithm demonstrates the best repeatability. Over 100 experiments, it achieved the lowest average fitness value of 84.66, with a standard deviation of 2.73. This represents a reduction of 5.5% in average fitness and 37.2% in standard deviation compared to HWPSO, which had values of 89.55 and 4.35, respectively. Furthermore, IPSO shows reductions of 23.7% in average fitness and 74.3% in standard deviation compared to PSO. It also outperformed WOA, achieving reductions of 24.6% in average fitness and 88.8% in standard deviation.

To demonstrate the effectiveness of our fitting method, we performed biarc curve fitting on the generated point sets shown in Fig. 5 (a), focusing on the case of two arcs and four arcs for each path. The curvatures of fitting arcs range from 0 to 0.02 (1/mm), which can be experimentally measured. We defined the direction vectors at the target and initial points as the unit vectors derived from the difference between two adjacent points, which helps align with the trajectory. The ratio ρ was varied from 0.1 to 3, with a step size of 0.1. Our optimization objective was to minimize the sum of distances between the fitted multi-arc trajectory and the original path (referred to as fit deviation). In Fig. 6, we present the biarc curves fitted for each algorithm, alongside the corresponding light-colored point sets depicting the original trajectories. The relevant data is presented in TABLE III. Among the fitted trajectories, the one obtained from the IPSO algorithms demonstrated the best performance in terms of smoothness, biarc length and deviation from the original path, with values of 0.27, 80.61 and 64.85 respectively. In Fig. 6 (b), it was observed that increasing the number of arc segments used for fitting significantly reduces the deviation between the fitted path and the initial path. The fit deviation values are 94.63, 78.04, 114.14, and 32.02 for WOA, PSO, HWPSO and IPSO, respectively. Reductions of 83.7%, 85.1%, 34.4%, and 50.6% were observed compared to the biarc fitting data. This adjustment allows the fitted path to better align with the initial path. However, increasing the number of arc segments may lead to a trajectory that is not feasible for practical puncture procedures. Considering the bending capability and minimum curvature of the flexible needle, we aim for a less complex fitted arc trajectory in upcoming puncture experiments.

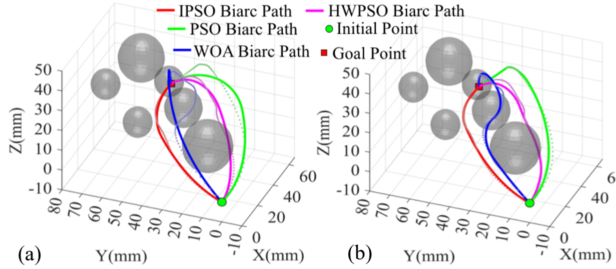

Fig. 6. The arc-based trajectories fitted from the discrete points generated by the path planning algorithms. (a) Case 1: Biarc path fitting. (b) Case 2: Four-arc (two biarcs) path fitting.

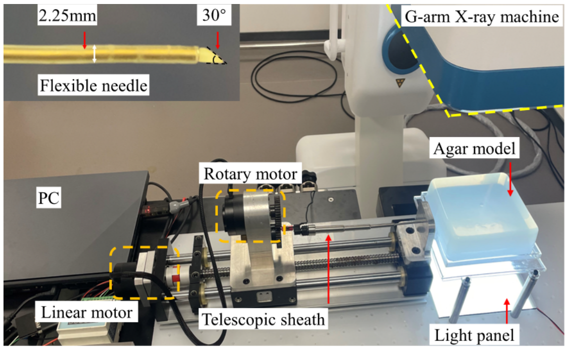

Fig. 7. Experimental platform for the flexible needle puncture.

TABLE II. COMPARISON RESULTS WITH SINGLE/MULTIPLE OBSTACLES

Algorithm	Max	Min	Mean	S_{td}
WOA	93.96 / 222.73	76.66 / 91.55	78.23 / 112.22	2.43 / 24.47
PSO	102.04 / 182.31	80.32 / 98.11	87.15 / 110.90	4.15 / 10.61
HWPSO	77.67 / 101.38	76.59 / 82.09	76.94 / 89.55	0.25 / 4.35
IPSO	**77.12 / 90.69**	**76.55 / 81.03**	**76.61 / 84.66**	**0.10 / 2.73**

TABLE III. PATH FITTING WITH TWO/FOUR ARCS

Algorithm	μ	Best ρ	Biarc length	Fit deviation
WOA	1.10	0.6 / [1.1,1.9]	90.50 / 95.31	582.31 / 94.63
PSO	0.88	1.1 / [0.1,3.0]	97.68 / 99.97	523.52 / 78.04
HWPSO	0.66	0.7 / [1.1,3.0]	85.75 / 87.39	174.01 / 114.14
IPSO	**0.27**	**0.6 / [0.1,0.5]**	**80.61 / 80.85**	**64.85 / 32.02**

B. Needle Puncture Experiments

We developed an experimental platform for flexible needle puncture, which primarily consists of a two-degree-of-freedom puncture mechanism with a telescopic sheath. The puncture trajectory of the flexible needle can be visually monitored by activating the light panel. High-precision imaging was achieved using a G-arm X-ray machine (Hiwise Medical) to obtain the needle position. The needle used in the experiment is a custom prototype made of a nitinol spiral coil (Outer diameter: 1.9 mm, Inner diameter: 1.6 mm, Length: 250 mm and Spring constant: 0.80 N/m). After being covered with polyolefin heat-shrink tube, the outer diameter measures 2.25 mm. The needle tip features a 30-degree 3D-printed cap in resin material. The agarose (BioFroxx) at a concentration of 0.6% mixed with distilled water was used to simulate brain tissue for the puncture experiments [26]. The components used for flexible needle puncture experiment are shown in Fig. 7.

Duty-cycling enables modulation of the needle curvature. To characterize the relationship between the duty cycle and curvature, numerous puncture experiments were performed. The feed speed was kept constant throughout the entire process. During each duty cycle, the needle rotates uniformly until the end of that cycle, maintaining a constant rotation speed during the rotation time τ_{rot}. The needle then stops rotating in the phase of τ_{stop}. In this experiment, the puncture speed $v(t)$ controlled by linear servo motor (PDA01, Daran Tech, China) is set at 5mm/s. Meanwhile, the rotation speed $\omega(t)$ of rotary motor is configured to π rad/s.

In the duty cycle validation experiment, each puncture procedure was conducted for 20 seconds. The flexible needle features a puncture length of 10 cm, with its initial tip direction perpendicular to the agarose gel. Puncture was repeated for 10 times for each duty cycle ranging from 0% to 100% in 25% increments. X-ray imaging was performed to capture the needle's trajectory, from which pixel coordinates were extracted. The pixel size was converted to actual

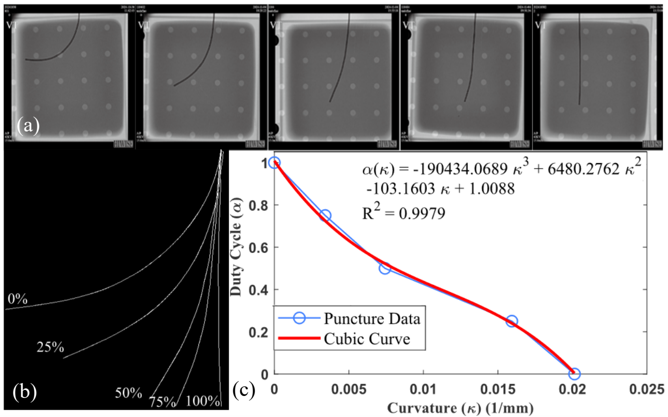

Fig. 8. (a) Pictures captured from the X-ray machine under different duty cycles. (b) Needle shape extraction with $\alpha = 0, 0.25, 0.5, 0.75$ and 1; (c) Duty cycle-curvature relationship fitting with cubic polynomial curve.

TABLE IV. DUTY CYCLE VALIDATION EXPERIMENTAL DATA

Duty-cycle	Average radius (mm)	S_{td}	Curvature (mm^{-1})
0%	49.65	4.25	0.0201
25%	62.76	2.85	0.0159
50%	134.72	18.80	0.0074
75%	293.35	51.52	0.0034
100%	+∞	0	0

physical dimensions using a calibration board. Then, arc fitting was applied using the least squares method to determine the bending radius. The results for the radius, curvature, and standard deviation for each duty cycle are summarized in TABLE IV. As shown in Fig. 8 (c), the relationship between duty cycle (denoted by α) and curvature (denoted by κ) can be approximated by a cubic curve:

$$\alpha(\kappa) = -1.90 \cdot 10^5 \kappa^3 + 6.48 \cdot 10^3 \kappa^2 - 103.16\kappa + 1.01 \quad (18)$$

which is consistent with the results presented in prior studies [27], with the corresponding X-ray images extracted in Fig. 8 (a), and (b). The curve exhibits a shape that is initially convex and then concave. As the duty cycle increases, the curvature continuously decreases, with the radius increases accordingly. From this one-to-one mapping curve, we determined the maximum curvature of 0.0201 mm^{-1} for our needle-tissue phantom, with the corresponding minimum radius being 49.65 mm, when duty cycle equals 0.

After characterizing the relationship between curvature and duty cycle, Puncture environments were designed with multiple obstacles. An optimal path was planned using the IPSO and biarc fitting algorithms as outlined in Section III with the values of N, c_3, m, λ, and ρ to 5, 0.095, 150, 100 and 1.5. The curvature constraint ranges from 0 to 0.0201 mm^{-1}. Based on the fitted curvatures, the matching duty cycle sequences were determined for the puncture control. The workspace size was set to 100 mm × 100 mm × 50 mm, with the start point and target point marked by red dots in Fig. 9 - Fig. 11.

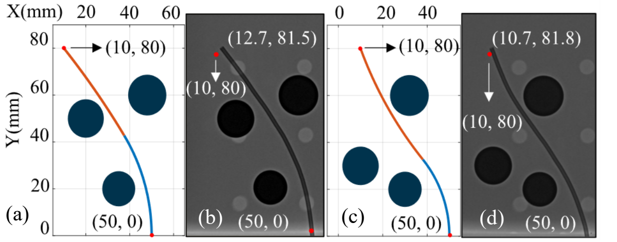

Fig. 9. Needle paths planned in C-shape (a) and S-shape (c). Needle trajectory captured by X-ray machine: (b) C-shape (d) S-shape.

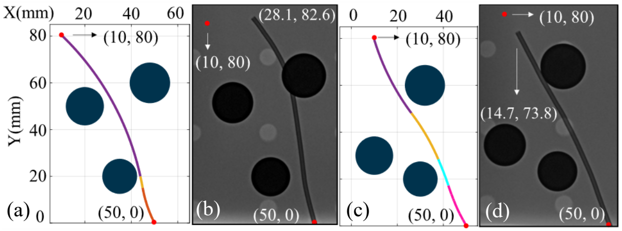

Fig. 10. (a)&(c) Paths planned by MT-RRT in [12]. (b)&(d) Needle trajectories captured by X-ray machine.

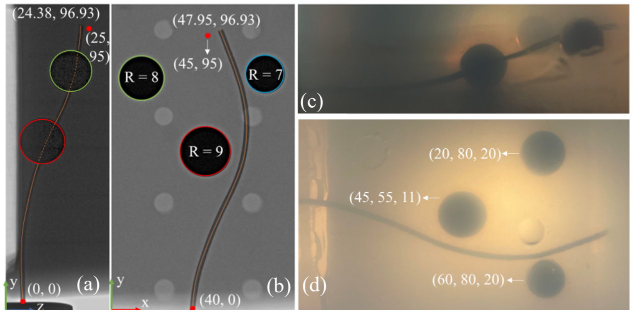

Fig. 11. 3D needle path captured by X-ray imaging in dual planes: (a) z-y plane, (b) x-y plane, and by camera from two views: (c) lateral (d) vertical.

A C-shape trajectory was obtained in the first experiment, with sphere obstacles of radii 8.2, 8.7, and 7.5 mm placed at (20, 50), (48, 60), and (35, 20), respectively, as shown in Fig. 9 (a). The radius and length of the proximal arc were 79.25 and 45.14 mm, while the distal arc had a radius and length of 351.15 and 46.34 mm, respectively. The corresponding duty cycles for the two arcs were 36% and 76%. After the insertion, the C-shape path was captured through X-ray imaging, illustrated in Fig. 9 (b). The actual position of the needle tip at the end of puncture was (12.7, 81.5), which is 3.09 mm away from the target point. An S-shape trajectory was presented in the second experiment, with obstacles of radii 8.2, 8.7, and 7.5 mm positioned at (10, 30), (32, 60), and (30, 20), respectively, as showed in Fig. 9 (c). The radius and length of the first arc were 50.74 and 35.60 mm, respectively, while the second arc measured 165.90 in radius and 55.19 mm in length. The duty cycles for two arcs were 3.4% and 58%. Fig. 9 (d) shows the X-ray image of the S-shape path captured after completing the puncture. The final position of the needle tip was (10.7, 81.8), resulting in a targeting error of 1.93 mm. Additionally, 10 repeated trials were conducted in both C-shape and S-shape environments, achieving a mean position error of 2.02 ± 0.66 mm and 3.01 ± 0.59 mm, respectively.

Fig. 10 (a) and (c) depict the two paths generated by the MT-RRT algorithm [12] under the identical environments defined in Fig. 9. The first path comprises three arcs with a total length of 92.28 mm, while the second path consists of four arcs totaling 92.9 mm. In contrast, IPSO produced shorter paths (91.48 mm and 90.79 mm respectively) with fewer arcs. More short arcs can cause significant deviation due to increased torsional windup resulting from frequent duty cycle adjustment. Across 10 repeated trials for each MT-RRT path, we obtained mean positioning errors of 25.41 ± 5.62 mm and 15.69 ± 5.36 mm, alongside collisions.

Fig. 11 illustrates a 3D needle steering experiment with a dual-arc S-shaped trajectory planned by our method. The first arc, driven by a 28% duty cycle, had a radius of 67.95 mm and a length of 44.75 mm, whereas the second arc (operated at 0% duty cycle) featured a reduced radius of 49.75 mm and extended 47.34 mm. To achieve orientation shifts between the 3D bending planes, angular adjustments of 33.5° and 26.7° were applied. Across 10 repeated trials, the system achieved a mean positioning error of 3.38 ± 0.37 mm, demonstrating consistent accuracy. In summary, the final puncture errors are within 4 mm, which is considered acceptable [8]. The errors may arise from the tissue deformation upon needle-tissue interaction, and the rotational lag between the needle tip and base caused by unexpected torsional movements.

V. CONCLUSION AND FUTURE WORK

In this study, we propose a novel approach that combines a good point set and a hybrid mutation technique to enhance the PSO-based path planning algorithm. Our aim is to improve its global search capability while achieving a high convergence speed. Additionally, 3D biarc curve fitting is employed to align the planned trajectory with the kinematics of the bevel-tip flexible needle. A comprehensive parametric study under controlled variable conditions was conducted to provide guidance on selecting the optimal parameters for the IPSO algorithm. Comparative analysis was conducted through numerous simulation tests in both single-obstacle and multi-obstacle environments, demonstrating that the proposed algorithm achieved optimal stability and superior solution sets. We also investigated the effectiveness of the biarc curve fitting method. Prior to conducting actual puncture experiments, we performed multiple trials to establish the relationship between duty cycles and curvatures for the specific interaction between the pair of flexible needle and tissue. This duty cycle-curvature mapping was subsequently applied in 2D and 3D experiments, validating the feasibility of the proposed algorithm. Our research demonstrates the capability of flexible needle to access region of interest while avoiding critical structures, ensuring an optimized curved path. We believe that with further improvements in control accuracy, this approach could have significant potential for future clinical applications, such as biopsy and deep brain stimulation.

In future research, closed-loop control of needle puncture trajectory using real time feedback from X-ray images or cameras will be investigated. Dynamic compensation on the rotation lag of needle shaft will be integrated to correct deviations in the trajectory. Intraoperative path replanning will also be studied to account for the shift of obstacles and target.

REFERENCES

[1] M. Li, D. Gao, Y. Lei, and T. Xu, "Dynamic path planning for bevel-tip flexible needle insertion into soft tissue based on a real-time finite element model," *Math. Probl. Eng.*, vol. 2020, pp. 1-13. Jul. 2020.

[2] Y. Zhao *et al.*, "Path planning for flexible needle based on both insertion mechanism kinematics and needle bending model," *Physics in Medicine & Biology*. 2024.

[3] K. Wu, B. Li, Y. Zhang, and X. Dai, "Review of research on path planning and control methods of flexible steerable needle puncture robot," *Comput. Assisted Surg.*, vol. 27, no. 1, pp. 91-112, Dec. 2022.

[4] V. Duindam, J. Xu, R. Alterovitz, S. Sastry, and K. Goldberg, "3D motion planning algorithms for steerable needles using inverse kinematics," *Algorithmic Foundation of Robotics viii*, vol. 57, pp. 535-549, 2010.

[5] J. Lee and W. Park, "Insertion planning for steerable flexible needles reaching multiple planar targets," in IEEE/RSJ International Conference on Intelligent Robots and Systems, Tokyo, Japan, 2013, pp. 2377-2383.

[6] J. Wang, X. Li, J. Zheng, and D. Sun, "Dynamic path planning for inserting a steerable needle into a soft tissue," *IEEE/ASME Trans. Mechatron.*, vol. 19, no. 2, pp. 549-558, Apr. 2014.

[7] B. Huo *et al.*, "Closed-loop control of bevel-tip needles based on path planning," *Robotica*, vol. 36, no. 12, pp. 1857-1873-447, Jul. 2018.

[8] C. Cai, C. Sun, Y. Han, and Q. Zhang, "Clinical flexible needle puncture path planning based on particle swarm optimization," *Comput. Methods Programs Biomed.*, vol. 193, p. 105511, Sept. 2020.

[9] F. Wang, H. Zhang, and A. Zhou, "A particle swarm optimization algorithm for mixed-variable optimization problems," *Swarm Evol. Comput.*, vol. 60, pp. 100808, Feb. 2021.

[10] S. Patil and R. Alterovitz, "Interactive motion planning for steerable needles in 3D environments with obstacles," in 3rd IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics, Tokyo, Japan, 2010, pp. 893-899.

[11] M. C. Bernardes, B. V. Adorno, G. A. Borges, and P. Poignet, "3D robust online motion planning for steerable needles in dynamic workspaces using duty-cycled rotation," *J. Control, Autom. Electr. Syst.*, vol. 25, no. 2, pp. 216-227, Apr. 2014.

[12] A. N. Aghdam and P. X. Liu, "A novel path planner for steerable bevel-tip needles to reach multiple targets with obstacles," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 10, pp. 7636-7645, Oct. 2020.

[13] Y. Zhang, Z. Ju, H. Zhang, and Z. Qi, "3-D path planning using improved RRT* algorithm for robot-assisted flexible needle insertion in multilayer tissues," *IEEE Can. J. Electr. Comput. Eng.*, vol. 45, no. 1, pp. 50-62, Dec. 2022.

[14] Y. Huang and F. Zhang, "Variable curvature path planning for robot-assisted flexible needle insertion based on improved Bi-RRT algorithm," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1-14, Mar. 2024.

[15] Y. Huang, L. Yu, and F. Zhang, "A survey on puncture models and path planning algorithms of bevel-tipped flexible needles," *Heliyon*, vol. 10, no. 3, p. e25002, Feb. 2024.

[16] K. Jerg, L. Boggaram Naveen, G. Kanschat, E. Okonkwo, and J. Hesser, "Diffuse domain approach for flexible needle insertion and relaxation," *Int. J. Numer. Methods Biomed. Eng.*, vol. 40, pp. e3782, Jan. 2024.

[17] J. Deng, S. Liu, Y. Liu, L. Wang, X. Gao, and K. Li, "A 2-DOF needle insertion device using inertial piezoelectric actuator," *IEEE Trans. Ind. Electron.*, vol. 69, no. 4, pp. 3918-3927, 2022.

[18] C. Rossa, N. Usmani, R. Sloboda, and M. Tavakoli, "A hand-held assistant for semiautomated percutaneous needle steering," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 3, pp. 637-648, Mar. 2017.

[19] B. Zhao *et al.*, "Curve fitting-based dynamic path planning and tracking control for flexible needle insertion," *IEEE Trans. Med. Rob. Bionics*, vol. 4, no. 2, pp. 436-447, May, 2022.

[20] K. Halder and M. F. Orlando, "Needle steering controller design for flexible steerable needle utilizing robust backstepping control strategy," *IEEE Trans. Med. Rob. Bionics*, vol. 6, no. 3, pp. 1256-1269, 2024.

[21] A. Tiwari and K. C. Pati, "Optimal control, stability and numerical integration analysis of unicycle," *Int. J. Dyn. Control*, vol. 9, no. 3, pp. 1042-1052, Sept. 2021.

[22] F. Zhao, F. Ji, T. Xu, N. Zhu, and Jonrinaldi, "Hierarchical parallel search with automatic parameter configuration for particle swarm optimization," *Appl. Soft Comput.*, vol. 151, p. 111126, Jan. 2024.

[23] G. He and X. Lu, "Good point set and double attractors based-QPSO and application in portfolio with transaction fee and financing cost," *Expert Syst. Appl.*, vol. 209, pp. 118339, Dec. 2022.

[24] X. Yang, J. Yan, D. Wang, Y. Xu, and G. Hua, "WOAD3QN-RP: An intelligent routing protocol in wireless sensor networks — A swarm intelligence and deep reinforcement learning based approach," *Expert Syst. Appl.*, vol. 246, pp. 123089, Jul. 2024.

[25] N. M. Laskar *et al.*, "HWPSO: A new hybrid whale-particle swarm optimization algorithm and its application in electronic design optimization problems," *Appl. Intell.*, vol. 49, no. 1, pp. 265-291, Jan. 2019.

[26] P. Muresan *et al.*, "Development of nanoparticle loaded microneedles for drug delivery to a brain tumour resection site," *Eur. J. Pharm. Biopharm.*, vol. 182, pp. 53-61, Jan. 2023.

[27] Y. Zhao, J. Lu, Y. Zhang, H. Zhang, and Y. Pei, "Structural design and puncture control of insertion robot for FBG embedded cannula flexible needle," *IEEE Access*, vol. 11, pp. 79295-79304, 2023.