SeedPro: Place Brachytherapy Seeds like Expert Clinicians via Hierarchical Reinforcement Agents

Haitao Li[1], Jiaxuan Liu[1], Wei Huang[2], Zhongmin Wang[2], and Xiaojun Chen[1]**

[1] Shanghai Jiao Tong University, Shanghai 200240, China
[2] Ruijin Hospital Shanghai Jiao Tong University School of Medicine, Shanghai, China

Abstract. Brachytherapy is an evidence-based, well-established treatment for intracorporeal tumors. A high-quality preoperative planning scheme is critical to ensuring optimal intraoperative efficacy. However, existing preoperative planning approaches often suffer from poor reproducibility, suboptimal dose distributions, and labor-intensive workflows. To address these limitations, we propose *SeedPro*, an automatic and fine-grained preoperative planning framework for brachytherapy that efficiently produces expert-level treatment plans with reduced computational costs, driven by hierarchical reinforcement learning agents. Experiments show that *SeedPro* achieves 12–20% reduction in OAR exposure and reduces puncture numbers by 25% compared to SOTA on both internal and external datasets. This work establishes a practical paradigm for intelligent and precision-oriented brachytherapy. The source code is available at `https://github.com/Haitao-Lee/SeedPro`.

Keywords: Brachytherapy · Reinforcement learning · Surgical planning.

1 Introduction

Brachytherapy is an evidence-based, well-established treatment for intracorporeal tumors [1]. By directly placing radioactive seeds intratumourally, it delivers a potent radiation dose that effectively kills cells nearby [4], as shown in Fig. 1. Compared with external beam radiotherapy, brachytherapy provides highly precise radiation with a prolonged therapeutic effect and minimal impact on healthy tissues, making it increasingly important for tumor suppression [6].

Accurate preoperative planning plays a pivotal role in brachytherapy, as it determines the dosimetric quality and therapeutic efficacy [7]. The core task is to devise a precise seed placement strategy that defines the number of radioactive seeds, their injection trajectories, and the spatial position and orientation of each seed. The conventional planning process typically begins with segmentation of the clinical target volume (CTV) and organs at risk (OARs) from preoperative images. Subsequently, clinicians manually review image slices to iteratively add

* *Corresponding author. Email: xiaojunchen@sjtu.edu.cn

or refine seed placements based on their clinical expertise and feedback from dose-calculation results generated by the treatment planning system (TPS) [9]. Throughout this process, dosimetric performance metrics, such as dose-volume histograms (DVH) [28], are continuously evaluated to iteratively refine the seed placement strategy until the prescribed treatment objectives are achieved. Intraoperatively, for each trajectory, a puncture needle is injected into the lesion to the predefined depth guided by surgical navigation or patient-specific templates [24], and as it is gradually withdrawn, radioactive seeds are deposited at specific intervals—an approach reminiscent of "laying eggs" along the trajectory [25].

However, the aforementioned planning approach exhibits significant limitations, including poor reproducibility, limited planning granularity, suboptimal dose distribution, and labor-intensive, time-consuming procedures [8]. Rapid advances in AI have enabled fast and precise 3D dose calculation in brachytherapy based on imaging data and the spatial distribution of seeds [11]. Compared with conventional Monte Carlo simulations [15] in the mainstream TPSs, the computation time for a single seed has been reduced to the millisecond scale, enabling the computationally intensive inverse seed placement planning to be performed efficiently, thereby potentially addressing the aforementioned limitations.

Several pioneering studies have explored related strategies. The current SOTA solution is *BrachyPlan* [14], which still relies on rule-based searching, and may perform suboptimally when dealing with irregularly shaped tumors. [16] introduced BRIGHT, an AI-driven framework capable of generating multiple near-optimal plans, while clinical evaluation [17] showed limited performance and a time-intensive workflow. [18] presented a knowledge-based approach for cervical brachytherapy, enhancing plan consistency and treatment outcomes. [26] applied reinforcement learning to brachytherapy, producing plans of comparable or higher quality with fewer needles. Other relevant efforts include [22, 23, 19]. The above studies indicate that deep reinforcement learning has emerged as a valuable tool for brachytherapy planning. Nevertheless, these methods share a common limitation: they predominantly focus on optimizing dose distribution, often yielding plans with excellent dosimetric performance but mechanically rigid seed arrangements that lack clinical practicality. One manifestation of this issue is that, in pursuit of dosimetric metrics, the resulting plans may involve an excessive number of punctures, sometimes requiring a separate puncture for a single seed [17]. In contrast, expert manual planning also places particular emphasis on surgical complexity and intraoperative implantation efficiency [33]. Therefore, it is significant to develop a planning system that places brachytherapy seeds like expert clinicians while ensuring reproducibility, efficacy, and efficiency.

2 Methods

Against the above background, we propose *SeedPro*, a fine-grained brachytherapy planning framework that enables efficient and expert-like seed placement through hierarchical reinforcement agents, as shown in Fig. 1. The main contributions are: 1) A multi-level state space generation algorithm that emulates

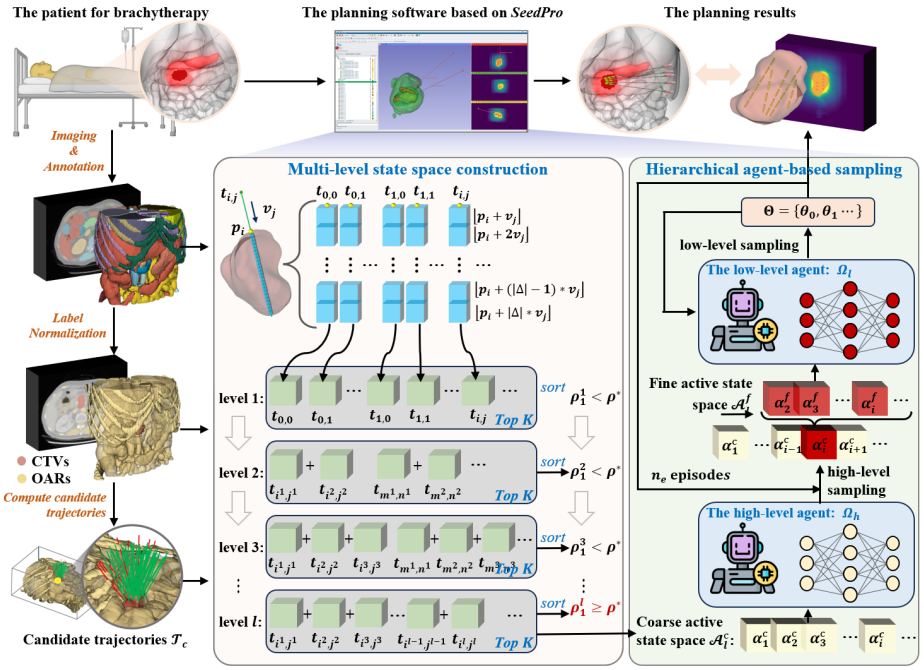

Fig. 1: The workflow of the *SeedPro*.

clinical decision-making patterns, enabling global fine-grained optimization; 2) A coarse-to-fine hierarchical reinforcement agents-based sampling strategy; 3) A segmented adaptive objective function ensuring compliance with CTV dosimetric endpoints while minimizing OARs exposure.

Expert-like multi-level state space construction: The inputs are preoperative images $\mathcal{I}_o \in \mathbb{R}^{D\times H\times W}$ and segmented masks $\mathcal{I}_s$ (with CTV, OARs, and background assigned values of ϵ_c, ϵ_o, and 0, respectively). From $\mathcal{I}_s$, the 3D coordinates of all CTV voxels are extracted to form a point cloud $\mathcal{P}_c = \{p_i\}_{i=1}^{|\mathcal{P}_c|}$. Following [14], the candidate trajectory set is constructed as $\tilde{\mathcal{T}}_c = \{t_{i,j} = \{\langle p_i, v_j\rangle, \Delta\}\}$, where p_i is the entry point, v_j is the direction, and Δ defines the permissible step sizes within $\mathcal{P}_c$. To avoid geometric interference, each trajectory $t_{i,j} \in \tilde{\mathcal{T}}_c$ is densely populated with seeds $S_{i,j}$, yielding the lowest-level active state space $\mathcal{A}_1 = \{\{\langle t_{i,j}, S_{i,j}\rangle\}, \cdots\}$. For higher levels $l = 2, 3, \ldots$, the corresponding active state space is recursively constructed as:

$$\begin{aligned} &\mathcal{A}_l = \{\{\langle t_{i_1,j_1}, S_{i_1,j_1}\rangle, \cdots, \langle t_{i_l,j_l}, S_{i_l,j_l}\rangle\}, \cdots\} \\ &s.t. \quad \{\langle t_{i_1,j_1}, S_{i_1,j_1}\rangle, \cdots, \langle t_{i_{l-1},j_{l-1}}, S_{i_{l-1},j_{l-1}}\rangle\} \in \mathcal{A}_{l-1}, \\ &\qquad \min_{n=1,\cdots,l-1} \|S_{i_l,j_l}, S_{i_n,j_n}\|_2 > 2r_s, \quad \{\langle t_{i_l,j_l}, S_{i_l,j_l}\rangle\} \in \mathcal{A}_1 \end{aligned} \tag{1}$$

where each element in the l-level active state space $\mathcal{A}_l$ consists of l non-interfering seed-filled trajectories, $\|\cdot\|_2$ represents the minimum Euclidean distance between seeds, and r_s is the seed radius.

During the construction of $\mathcal{A}_l$, the cumulative dose field $\Psi_m^l \in \mathbb{R}^{D\times H\times W}$ induced by all seeds in each element $\pi_m^l \in \mathcal{A}_l$ is computed. A dosimetric evaluation function calculates the corresponding metric $\rho_m^l = f_{\text{eval}}(\mathcal{I}_s, \Psi_m^l, \epsilon_c, \cdots)$. To reduce computational complexity, a top-K selection strategy is applied: elements are ranked in descending order by ρ_m^l, and only the top K configurations are retained. The state space is then updated to the filtered set $\mathcal{A}_l = \{\pi_1^l, \cdots, \pi_k^l\}$.

Given a prescribed dosimetric objective ρ^*, if the highest metric $\rho_1^l < \rho^*$, none of the trajectory combinations satisfy the requirement, and the process is repeated to construct $\mathcal{A}_{l+1}$. Conversely, if $\rho_1^l \geq \rho^*$, a feasible trajectory combination emerges at level l. Following the expert-style paradigm—which prioritizes minimizing puncture needles over reducing implanted seeds—subsequent optimization is confined to $\mathcal{A}_l$. Hierarchically built from previous levels, $\mathcal{A}_l$ serves as the active domain for reinforcement agents, enhancing clinical consistency while reducing search dimensionality and improving planning efficiency.

Hierarchical reinforcement learning-based sampling: For $\pi_k^l \in \mathcal{A}_l$, each available position in its trajectories serves as a potential state. Let $\Theta = \{\theta_i\}$ be the initially empty set of placed seeds, where $\theta_i = \{\langle p_i, v_i\rangle, \delta_i\}$ indicates placement at $\lfloor p_i + \delta_i v_i \rfloor$ with orientation v_i. Using a distance map $\mathcal{I}_d$ of the CTV, feasible positions for new seeds along each trajectory in $\mathcal{A}_l$ are calculated:

$$\begin{aligned} &\alpha_k^l = \left\{\delta \mid \forall\delta \in \Delta, \forall\langle t_{i,j} = \{\langle p_i, v_j\rangle, \Delta\}, S_{i,j}\rangle \in \pi_k^l\right\} \\ &s.t. \quad \min_{\forall\theta_m\in\Theta} \|(\delta v_j - \delta_m v_m)v_j\|_2 > l_s \vee v_m \neq v_j \vee p_m \neq p_j, \;\; \mathcal{I}_d[\lfloor p_i + \delta v_j \rfloor] \geq \frac{l_s}{2} \end{aligned} \tag{2}$$

where $\lfloor\cdot\rfloor$ is the floor function and l_s is the seed length. The flattened, coarse active state space is defined as $\mathcal{A}_l^c = \bigcup_{k=1}^{|\mathcal{A}_l|} \alpha_k^l = \{\alpha_1^c, \alpha_2^c, \dots\}$, where α_i^c is a candidate action.

A hierarchical policy on $\mathcal{A}_l^c$ efficiently identifies optimal plans. The high-level agent Ω_h coarsely selects needle trajectory combinations from $\mathcal{A}_l$, while the low-level agent Ω_l determines the optimal positions of individual seeds along each selected trajectory until dosimetric objectives are met. Both use lightweight MLPs: Ω_h outputs $(b, |\mathcal{A}_l^c|)$ logits, and Ω_l adapts its output dimension dynamically. Initialized at state $\omega_c = 0$, Ω_h generates $z_h \in \mathbb{R}^{b\times|\mathcal{A}_l^c|}$ to sample a high-level action $\delta_h \sim p(\alpha_i^c \mid \omega_c) = \frac{\exp(z_{h,i})}{\sum_j \exp(z_{h,j})}$. Using δ_h's index i_h, the trajectory and subset indices (n^*, k^*) are determined by $\sum_{n=1}^{k-1}|\alpha_n^l| + \sum_{m=1}^{n-1}|\Delta_m| \leq i_h < \sum_{n=1}^{k-1}|\alpha_n^l| + \sum_{m=1}^{n}|\Delta_m|$ for $\pi_k^l \in \mathcal{A}_l$. The trajectory for the first seed is $t_{i_{n^*},j_{n^*}} \in \pi_{k^*}^l$. We add $\theta_0 = \{\langle p_{i_{n^*}}, v_{j_{n^*}}\rangle, \delta_h\}$ to Θ, construct the fine active state

space $\mathcal{A}_l^f = \alpha_{k^*}^f$, and generate an availability mask $\mathcal{M}$ to prevent interference:

$$\begin{aligned} &\mathcal{M} = \{m_1, m_2, \cdots, m_{|\mathcal{A}_l^f|}\} \\ s.t. \quad &m_i = \mathbb{I}\left(\min_{\forall \theta_j \in \Theta} \|(\alpha_i^f v_{i_h} - \delta_j v_j) v_i\|_2 > l_s \vee v_{i_h} \neq v_j \vee p_{i_h} \neq p_j\right) \\ &\text{for} \ \ \textstyle\sum_{m=1}^{h-1} |\Delta_m| \leq i < \sum_{m=1}^{h} |\Delta_m|, \ \ h = 1, 2, \cdots, l \end{aligned} \tag{3}$$

where $\mathbb{I}(\cdot)$ is the indicator function ($m_i = 1$ if feasible, 0 otherwise). Resetting $\omega_c = 0$, Ω_l produces a sampling vector $z_l \in \mathbb{R}^{b \times |\mathcal{A}_l^f|}$, converted to a masked probability $p(\alpha_i^f \mid \omega_c, \mathcal{M}) = \frac{\exp(z_{l,i})}{\sum_j \exp(z_{l,j})} \times m_i$. The next action $\delta_{l,1}$ is sampled, and $\theta_1 = \{\langle p_{l,1}, v_{l,1} \rangle, \delta_{l,1}\}$ is appended to Θ. This iteratively generates $\theta_2, \theta_3, \ldots$ until reaching the dosimetric endpoint or $\sum m_i = 0$, forming a complete solution.

This procedure runs per episode, with Ω_h constructing a dedicated policy network for Ω_l. Only Ω_h updates via backpropagation to approximate expert decisions, while Ω_l reinitializes. A two-stage learning strategy is employed: during the first $n_e/2$ episodes, agents explore global configurations. Addressing "trajectory under-saturation", the second stage refines the best scheme Θ^* and its fine state space $\mathcal{A}_l^{f,*}$. For the remaining $n_e/2$ episodes, Θ^* is cleared, and placement proceeds within $\mathcal{A}_l^{f,*}$ until targets are met, continuously updating Ω_l via backpropagation. Finally, the highest-performing trajectory from $\mathcal{A}_l$ serves as a reference to evaluate the final plan.

Segmented adaptive objective function: For the currently placed seeds $\Theta_c = \{\theta_0, \theta_1, \ldots, \theta_{n_c}\}$, where n_c denotes the number of seeds, the cumulative dose distribution Ψ_c is computed via the dose calculator. A predefined D_X threshold μ_X is specified, where X represents the percentage of the CTV volume required to receive at least μ_X Gy. Notably, the objective function is compatible with standard clinical metrics, including V_x, $D_{2\text{cc}}$, $D_{1\text{cc}}$, $D_{0.5\text{cc}}$, D_{mean}, or any institution-specific criterion, without requiring architectural modifications. Given these parameters, the objective function is formulated as follows:

$$\begin{aligned} &r_{n_c} = f_{eval}(\Theta_c, \mathcal{I}_o, \mathcal{I}_s, \mu_X, X) = \rho_{n_c} + \mathbb{I}(\rho_{n_c} \geq X) \, \tfrac{D_c}{D_o}, \\ &\text{s.t.} \begin{cases} \Psi_c = f_{dose}(\Theta_c, \mathcal{I}_o) \\ D_c = \displaystyle\sum_{p \in \mathcal{I}_s, \, \mathcal{I}_s[p] = \epsilon_c} \Psi_c[p], \\ D_o = \displaystyle\sum_{p \in \mathcal{I}_s, \, \mathcal{I}_s[p] = \epsilon_o} \Psi_c[p], \\ \rho_{n_c} = \dfrac{\sum_{p \in \mathcal{I}_s} \mathbb{I}((\mathcal{I}_s[p] = \epsilon_c) \wedge (\Psi_c[p] \geq \mu_X))}{\sum_{p \in \mathcal{I}_s} \mathbb{I}(\mathcal{I}_s[p] = \epsilon_c)} \times 100. \end{cases} \end{aligned} \tag{4}$$

where f_{dose} denotes the dose calculation function based on the dose calculator; $\mathbb{I}(a)$ denotes the indicator function, which returns 1 when the condition a is satisfied and 0 otherwise; $\wedge$ represents the logical AND operation, indicating the conjunction of multiple conditions in formal logic; $A[b]$ denotes the value of matrix A at coordinate b; r_{n_c} represents the reward obtained after placing the n_c^{th} seed θ_{n_c}, and the corresponding reward set is denoted as $R_c = \{r_0, r_1, \cdots, r_{n_c}\}$.

3 Experiments and Discussion

3.1 Dataset and experimental setting

To evaluate the proposed *SeedPro* system, we constructed an in-house clinical dataset of 17 patients from The Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, including cases of pancreatic, head and neck, and other tumor types, alongside 282 pancreatic cancer cases from the open-source external Medical Segmentation Decathlon (MSD) dataset [36]. All patients in our in-house dataset underwent ^{125}I seed brachytherapy, with each seed measuring 0.4 mm in radius and 4.5 mm in length. Prescribed CTV dosimetric objectives were set as $D_{90} \geq 120$ Gy and $V_{100} \geq 90\%$, following the OpenKBP standard [35]. CT scans were acquired with 5 mm slice thickness and 1 mm in-plane resolution, comprising 12–90 slices per volume depending on lesion size. CTVs and OARs were manually delineated by experienced interventional radiologists. Experiments were performed on a workstation with dual NVIDIA RTX 3090 GPUs, and the policy network was optimized with a learning rate of 10^{-3}. It is worth noting that sensitivity and ablation studies as well as robustness evaluations were conducted exclusively on the in-house dataset. The external dataset was specifically utilized in the comparative study to rigorously assess the generalizability of our method across a large cohort. Importantly, *SeedPro*, developed and validated on this small-scale in-house cohort, achieves superior performance on a substantially larger independent external dataset without any targeted tuning, directly demonstrating the robustness and reliability of the proposed framework.

3.2 Sensitivity study

Evaluation of K: In state space construction, a top-K selection strategy is introduced to reduce computational overhead in constructing the expert-inspired multi-level active state space. To evaluate its effect, K was set to 10, 50, 100, 200, and 400 to generate $\mathcal{A}_l$ according to Eq. (1), and the $\langle \pi_m^l, \Psi_m^l, \rho_m^l \rangle$ with the highest D_{90} was selected for evaluation. While exceeding the prescribed dosimetric targets does not imply better clinical quality, larger D_{90} and V_{100} here indicate greater optimization potential and higher-quality $\mathcal{A}_l$. The construction time for $\mathcal{A}_l$ was also recorded to quantify computational cost. Results are summarized in Fig. 2a, with SR denoting the success rate of meeting the prescribed V_{100} and D_{90}.

Evaluation of n_e: In the proposed coarse-to-fine two-stage learning strategy, the total number of episodes is defined as n_e. To assess the sensitivity of *SeedPro*'s performance with respect to different n_e settings, n_e was set to 50, 100, 200, 400, and 800 for comparative evaluation. To clearly demonstrate the direct influence of n_e on the reinforcement agents Ω_h and Ω_l, we did not employ the baseline mechanism introduced in Section 2; instead, the final results were generated solely by the reinforcement agents. We calculated the success rate (SR), V_{100}, D_{90}, and computation time for each configuration, as shown in Fig. 2b.

Table 1: Ablation study

Configuration	SR (%, ↑)$_{V_{100}(\%,\geq 90),D_{90}(Gy,\geq 120)}$	$\tilde{V}_{150}$ (%, ↓)	$\tilde{V}_{100}$ (%, ↓)	$\tilde{V}_{50}$ (%, ↓)	Times (s, ↓)
Baseline	100.00 $_{94.90\pm2.94,137.04\pm12.20}$	45.99 ± 31.44	150.10 ± 86.42	650.37 ± 303.48	**34.98±19.92**
+ segmented adaptive objective function	100.00 $_{90.76\pm0.80,121.81\pm1.82}$	28.30 ± 25.89	85.37 ± 54.87	502.90 ± 227.75	40.52±24.03
+ two stage learning strategy	100.00 $_{90.53\pm0.48,121.40\pm1.24}$	**27.78 ± 24.66**	**84.77 ± 53.97**	**502.85 ± 223.50**	37.30±21.46

Table 2: Quantitative comparison of different methods under the dosimetric requirement of $\mu_{90} = 120$ Gy, evaluated on both the internal (17 cases) and external (282 cases) datasets.

Dataset	Methods	SR (%, ↑)$_{V_{100}(\%,\geq 90),D_{90}(Gy,\geq \mu_{90})}$	$\tilde{V}_{150}$ (%, ↓)	$\tilde{V}_{100}$ (%, ↓)	$\tilde{V}_{50}$ (%, ↓)	n_p ↓	n_s ↓
Internal	BrachyPlan	100.00 $_{91.54\pm1.32,124.78\pm4.44}$	39.83 ± 24.27	130.80 ± 67.22	674.43 ± 351.97	8.06 ± 3.98	35.06 ± 19.72
	SeedPro (ours)	100.00 $_{90.53\pm0.48,121.40\pm1.24}$	**27.78 ± 24.66**	**84.77 ± 53.97**	**502.85 ± 223.50**	**6.06 ± 3.02**	**30.59 ± 18.59**
External	BrachyPlan	81.91 $_{93.20\pm5.15,141.22\pm22.71}$	69.67 ± 90.55	163.22 ± 370.01	670.07 ± 767.45	2.85 ± 1.01	8.62 ± 7.05
	SeedPro (ours)	**88.65** $_{91.25\pm4.33,125.51\pm13.86}$	**30.28 ± 92.24**	**86.18 ± 370.46**	**356.10 ± 755.19**	**2.45 ± 0.80**	**6.99 ± 6.89**

Table 3: Robustness study under $\mu_{90} = 100$ Gy and 140 Gy to simulate varying clinical scenarios.

	Methods	SR (%, ↑)$_{V_{100}(\%,\geq 90),D_{90}(Gy,\geq \mu_{90})}$	$\tilde{V}_{150}$ (%, ↓)	$\tilde{V}_{100}$ (%, ↓)	$\tilde{V}_{50}$ (%, ↓)	n_p ↓	n_s ↓
μ_{90} = 100 Gy	BrachyPlan	100.00 $_{92.26\pm1.90,106.55\pm5.81}$	73.17 ± 42.21	215.11 ± 104.48	804.30 ± 390.26	6.29 ± 3.34	31.94 ± 18.34
	SeedPro (ours)	100.00 $_{90.91\pm0.78,101.75\pm1.60}$	**32.95 ± 27.43**	**98.72 ± 58.94**	**573.66 ± 280.05**	**5.35 ± 2.81**	**26.12 ± 15.53**
μ_{90} = 140 Gy	BrachyPlan	94.12 $_{91.45\pm1.12,145.12\pm3.67}$	53.91 ± 28.19	152.11 ± 60.42	690.40 ± 309.06	8.50 ± 4.08	40.65 ± 24.03
	SeedPro (ours)	**100.00** $_{90.37\pm0.36,141.25\pm1.36}$	**27.31 ± 22.77**	**82.13 ± 52.26**	**475.26 ± 206.53**	**6.82 ± 3.50**	**34.76 ± 21.18**

3.3 Ablation study

We conducted an ablation study to evaluate the core modules of *SeedPro*. Setting the optimal trajectory combination from the multi-level state space as the baseline, we incrementally integrated the segmented adaptive objective function and the coarse-to-fine two-stage optimization strategy. Table 1 reports the success rate (SR), time cost, and $\tilde{V}_x$ (the fraction of OAR volume receiving $\geq x\% \times D_{90}$, where lower indicates better OAR sparing). Both modules progressively enhanced OAR protection while maintaining target coverage.

K	SR (%, ↑)	V_{100} (%, ≥90)	D_{90} (Gy, ≥120)	Times (s, ↓)
10	70.59	91.19±6.71	122.72±31.08	0.80±0.82
50	88.24	93.48±3.88	133.06±17.49	0.82±0.82
100	94.12	93.92±3.55	135.17±15.73	0.91±0.86
200	100.00	94.27±3.23	135.90±13.49	1.06±0.92
400	100.00	94.46±3.40	135.36±14.58	1.39±1.07

(a) Evaluation of K

n_e	SR (%, ↑)	V_{100} (%, ≥90)	D_{90} (Gy, ≥120)	Times (s, ↓)
50	64.71	89.84 ± 1.50	119.59±3.17	6.06±3.40
100	76.47	89.98 ± 1.55	119.45±3.38	10.70±6.26
200	76.47	90.01 ± 1.48	119.69±3.54	19.93±12.12
400	82.35	90.12 ± 1.32	120.17±3.48	37.30±21.46
800	82.35	90.29 ± 1.20	120.61±2.83	73.50±44.78

(b) Evaluation of n_e

Fig. 2: (a) Evaluation of K (b) Evaluation of n_e

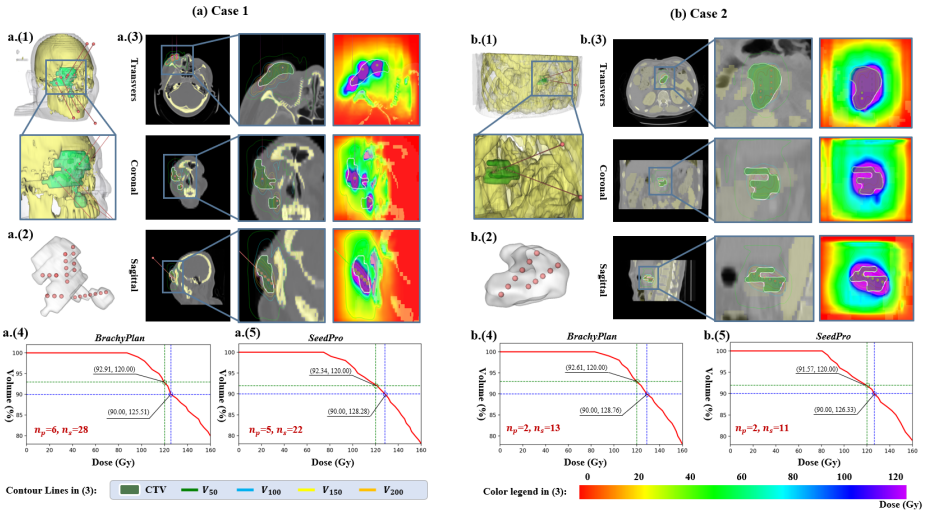

Fig. 3: The planning results of two randomly selected cases by *SeedPro*: (1-2) 3D visualization results; (3) dose distribution results with 2D CTV contour lines; (4-5) corresponding DVH curves.

3.4 Comparison study

To evaluate the superiority and generalizability of *SeedPro* on both in-house and external datasets, we selected the recent SOTA method, *BrachyPlan* [14], as our baseline. As *SeedPro* represents a direct methodological advancement over *BrachyPlan*—which has been rigorously validated to achieve clinical equivalence compared to expert manual plans—this comparison provides a stringent and representative evaluation of our method's clinical potential.

The evaluation metrics included the success rate (SR), dosimetric indices $\tilde{V}_x$, the number of puncture needles (n_p), and the number of implanted seeds (n_s). Here, n_p reflects procedural complexity—fewer punctures indicate greater intra-operative convenience—while n_s represents the precision of dose field control, as achieving comparable dosimetric outcomes with fewer seeds implies a more fine-grained and efficient planning strategy. The quantitative results are summarized in Table 2 and visually illustrated in Fig. 3. *SeedPro* consistently outperformed *BrachyPlan* across both datasets, particularly demonstrating superior OAR protection and a notable reduction in surgical complexity.

3.5 Robustness study

To further assess robustness and confirm safety under extreme or diverse prescription contexts, *SeedPro* was evaluated under varying dosimetric requirements. Specifically, we simulated clinical scenarios with different tumor malignancy grades by setting target doses of $D_{90} = 100$ Gy and 140 Gy. All comparative results against the baseline are summarized in Table 3. *SeedPro* maintained

exceptional consistency and robustness, achieving a 100% success rate across the diverse target doses while continuing to guarantee optimal tissue protection and high needle efficiency.

4 Conclusion

In this work, we propose *SeedPro*, a novel framework for expert-like brachytherapy planning, and demonstrate its superiority through comprehensive experiments. Future work will focus on incorporating larger-scale and multi-center clinical datasets to further enhance robustness and generalizability, and on validating *SeedPro* across additional anatomical sites, including prostate.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Chargari, C., Deutsch, E., Blanchard, P., et al.: Brachytherapy: An overview for clinicians. CA: A Cancer Journal for Clinicians **69**(5), 386–401 (2019)
2. Lim, Y.K., Kim, D.: Brachytherapy: A comprehensive review. Progress in Medical Physics **32**(2), 25–39 (2021)
3. Ragde, H., Grado, G.L., Nadir, B., et al.: Modern prostate brachytherapy. CA: A Cancer Journal for Clinicians **50**(6), 380–393 (2000)
4. Pathak, P., Thomas, J.J., Baghwala, A., et al.: Personalized brachytherapy: applications and future directions. Cancers **16**(19), 3424 (2024)
5. Pötter, R., Tanderup, K., Schmid, M.P., et al.: MRI-guided adaptive brachytherapy in locally advanced cervical cancer (EMBRACE-I): a multicentre prospective cohort study. The Lancet Oncology **22**(4), 538–547 (2021)
6. Serre, R., Gabro, A., Andraud, M., et al.: Brachytherapy: Perspectives for combined treatments with immunotherapy. Clinical and Translational Radiation Oncology, 100924 (2025)
7. Blasko, J.C.: Low-dose-rate brachytherapy for prostate cancer: Preplanning vs. intraoperative planning—Preplanning is best. Brachytherapy **5**(3), 139–142 (2006)
8. Matzkin, H., Chen, J., German, L., et al.: Comparison between preoperative and real-time intraoperative planning 125I permanent prostate brachytherapy: Long-term clinical biochemical outcome. Radiation Oncology **8**(1), 288 (2013)
9. Berger, D., Van Dyk, S., Beaulieu, L., et al.: Modern tools for modern brachytherapy. Clinical Oncology **35**(8), e453–e468 (2023)
10. Liu, T., Wen, S., Wang, S., et al.: Artificial intelligence in brachytherapy. Journal of Radiation Research and Applied Sciences **17**(2), 100925 (2024)
11. Chen, J., Qiu, R.L.J., Wang, T., et al.: A review of artificial intelligence in brachytherapy. Journal of Applied Clinical Medical Physics, e70034 (2025)
12. Gautam, S., Osman, A.F.I., Richeson, D., et al.: Attention 3D U–NET for dose distribution prediction of high-dose-rate brachytherapy of cervical cancer: direction modulated brachytherapy tandem applicator. Med. Phys. **51**(8), 5593–5603 (2024)
13. Quetin, S., Bahoric, B., Maleki, F., et al.: Deep learning for high-resolution dose prediction in high dose rate brachytherapy for breast cancer treatment. Phys. Med. Biol. **69**(10), 105011 (2024)

14. Liu, J., Li, H., Shi, H., et al.: BrachyPlan: A Fine-grained Efficient Dose-guided Inverse Planning Strategy for Low-dose-rate Brachytherapy. Medical Image Analysis, 103852 (2025)
15. Bonenfant, É., Magnoux, V., Hissoiny, S., et al.: Fast GPU-based Monte Carlo simulations for LDR prostate brachytherapy. Phys. Med. Biol. **60**(13), 4973 (2015)
16. Dickhoff, L.R.M., Scholman, R.J., Barten, D.L.J., et al.: Keeping your best options open with AI-based treatment planning in prostate and cervix brachytherapy. Brachytherapy **23**(2), 188–198 (2024)
17. Barten, D.L.J., Pieters, B.R., Bouter, A., et al.: Towards artificial intelligence-based automated treatment planning in clinical practice: A prospective study of the first clinical experiences in high-dose-rate prostate brachytherapy. Brachytherapy **22**(2), 279–289 (2023)
18. Gautam, S., Flower, E., Richeson, D., et al.: Improving plan quality in cervical cancer brachytherapy using knowledge-based planning for direction modulated brachytherapy tandem applicator. Brachytherapy **24**(1), 144–153 (2025)
19. Jiang, H., Li, A., Ke, K., et al.: Three-dimensional treatment-planning-based prediction of seed migration to chest after 125I seed brachytherapy for hepatic malignancy. Brachytherapy **23**(4), 478–488 (2024)
20. Pu, G., Jiang, S., Yang, Z., et al.: Deep reinforcement learning for treatment planning in high-dose-rate cervical brachytherapy. Phys. Med. **94**, 1–7 (2022)
21. Kallis, K., Moore, L.C., Cortes, K.G., et al.: Automated treatment planning framework for brachytherapy of cervical cancer using 3D dose predictions. Physics in Medicine & Biology **68**(8), 085011 (2023)
22. Xiao, Z., Xiong, T., Geng, L., et al.: Automatic planning for head and neck seed implant brachytherapy based on deep convolutional neural network dose engine. Medical Physics **51**(2), 1460–1473 (2024)
23. Aumüller, P.: An Inverse Treatment Planning System for Robotic Seed Brachytherapy. PhD thesis (2024)
24. Li, Y., Jiang, J., Jiang, Q., et al.: Advanced navigation technology enables endobronchial brachytherapy for peripheral lung cancer: an old technique plays a new role. Brachytherapy **23**(2), 199–206 (2024)
25. Hu, P., Huang, J., Zhang, Y., et al.: Iodine-125 seed implantation in the treatment of malignant tumors. Journal of Interventional Medicine **6**(3), 111–115 (2023)
26. Wang, T., Feng, Y., Yang, X.: Automatic Treatment Planning using Reinforcement Learning for High-dose-rate Prostate Brachytherapy. arXiv preprint arXiv:2506.09805 (2025)
27. Qian, Y., Washington, P., Podder, T.K., et al.: A Linear Programming and Deep Reinforcement Learning Framework to Choose Dwell Positions and Dwell Time in High-Dose-Rate Prostate Brachytherapy Using Curvilinear Catheters. In: ASME International Mechanical Engineering Congress and Exposition, vol. 88629, pp. V004T06A030. ASME (2024)
28. Kirisits, C., Siebert, F.A., Baltas, D., et al.: Accuracy of volume and DVH parameters determined with different brachytherapy treatment planning systems. Radiotherapy and Oncology **84**(3), 290–297 (2007)
29. Lindegaard, J.C., Tanderup, K., Nielsen, S.K., et al.: MRI-guided 3D optimization significantly improves DVH parameters of pulsed-dose-rate brachytherapy in locally advanced cervical cancer. International Journal of Radiation Oncology*Biology*Physics **71**, 756–764 (2008)
30. Merrick, G.S., Butler, W.M., Dorsey, A.T., et al.: Potential role of various dosimetric quality indicators in prostate brachytherapy. International Journal of Radiation Oncology, Biology, Physics **44**(3), 717–724 (1999)

31. Shi, J., Chen, J., He, G., et al.: Artificial intelligence in high-dose-rate brachytherapy treatment planning for cervical cancer: a review. Frontiers in Oncology **15**, 1507592 (2025)
32. Rossi, L., Bijman, R., Chopra, S., et al.: Rule-based AI automated adaptive treatment planning for image guided cervical cancer brachytherapy. Brachytherapy (2025)
33. Li, M., Li, X., Qiu, B., et al.: Experts consensus on 3D-printing template-assisted CT-guided radioactive iodine-125 seed implantation for recurrent soft tissue carcinoma in China. Clinical and Experimental Medicine **25**(1), 52 (2025)
34. Chitmanee, P., Sampaongen, T., Klomjit, N.: Factors affecting D90 high-risk clinical target volumes (hr-ctv dose) of intracavitary and interstitial brachytherapy in locally advanced cervical cancer. Clinical Oncology **37**, 103701 (2025)
35. Babier, A., Zhang, B., Mahmood, R., et al.: OpenKBP: The Open-Access Knowledge-Based Planning Grand Challenge and Dataset. Medical Physics **48**(9), 5549–5561 (2021)
36. Antonelli, M., Reinke, A., Bakas, S., et al.: The Medical Segmentation Decathlon. Nature Communications **13**, 4128 (2022)