Federated Deep Reinforcement Learning for Privacy-Preserving Robotic-Assisted Surgery

Sana Hafeez*, Sundas Rafat Mulkana, Muhammad Ali Imran†, and Michele Sevegnani*

*School of Computing Science, University of Glasgow, UK

†James Watt School of Engineering, University of Glasgow, UK

Emails: {sundasrafat.mulkana, Muhammad.Imran, Michele.Sevegnani}@glasgow.ac.uk

Abstract—The integration of Reinforcement Learning (RL) into robotic-assisted surgery (RAS) holds significant promise for advancing surgical precision, adaptability, and autonomous decision-making. However, the development of robust RL models in clinical settings is hindered by key challenges, including stringent patient data privacy regulations, limited access to diverse surgical datasets, and high procedural variability. To address these limitations, this paper presents a Federated Deep Reinforcement Learning (FDRL) framework that enables decentralised training of RL models across multiple healthcare institutions without exposing sensitive patient information. A central innovation of the proposed framework is its dynamic policy adaptation mechanism, which allows surgical robots to select and tailor patient-specific policies in real-time, thereby ensuring personalised and optimised interventions. To uphold rigorous privacy standards while facilitating collaborative learning, the FDRL framework incorporates secure aggregation, differential privacy, and homomorphic encryption techniques. Experimental results demonstrate a 60% reduction in privacy leakage compared to conventional methods, with surgical precision maintained within a 1.5% margin of a centralised baseline. This work establishes a foundational approach for adaptive, secure, and patient-centric AI-driven surgical robotics, offering a pathway toward clinical translation and scalable deployment across diverse healthcare environments.

Index Terms—Federated Deep Reinforcement Learning, Autonomous Surgical Robots, Task-Based Privacy-Preservation, Federated Learning, Differential Privacy, Secure Reinforcement Learning, Homomorphic Encryption, and Secure Aggregation.

I. INTRODUCTION

A. Background and Motivation

ROBOTIC-assisted surgery has revolutionised modern medicine, offering a paradigm shift from traditional open surgery to minimally invasive procedures. This transition has led to significant advancements, including enhanced surgical precision, diminished patient trauma, reduced post-operative complications, and accelerated recovery times [1]–[3]. Augmenting Robotic-Assisted Surgery (RAS) platforms with Artificial Intelligence (AI) is the next frontier, promising to further enhance surgical capabilities and autonomy. Specifically, the integration of AI can enable surgical robots to perform complex tasks with greater dexterity, adapt to unforeseen intraoperative events, and provide surgeons with real-time decision support [4]–[7].

Corresponding author: Sana Hafeez (email: sanahafeez2828@gmail.com).

Within the pantheon of AI methodologies, Reinforcement Learning (RL) has emerged as a particularly potent approach for endowing surgical robots with intelligent decision-making capabilities. RL empowers autonomous agents, in this case, surgical robots, to learn optimal sequences of actions through iterative interaction with a dynamic environment, guided by reward signals [5]. RL algorithms, when utilising real-time intraoperative information alongside historical procedural data, can empower surgical robots to enhance their control strategies, customise interventions to suit each patient's unique anatomical and physiological characteristics, and adjust in real-time to the unpredictable and variable nature of surgical procedures [6], [7]. This capability is crucial for navigating the nuanced and often complex landscape of surgical interventions.

B. Challenges in RL-based Surgical Robotics

Despite the transformative potential of RL in RAS, several formidable challenges impede its widespread adoption and clinical translation. A primary obstacle is the inherent heterogeneity of surgical environments. These environments are characterised by significant inter-patient anatomical variability, diverse patient comorbidities, and surgeon-specific procedural preferences, all of which contribute to a high degree of complexity and pose a substantial challenge to the generalisability of RL models [8]–[10]. Furthermore, the development of robust RL models necessitates access to large, diverse datasets of surgical procedures. However, individual healthcare institutions often suffer from data scarcity, which limits the ability of RL models trained in isolation to generalise effectively to real-world clinical settings, particularly when encountering rare pathologies or unanticipated procedural complexities [11].

Moreover, RL-based surgical systems rely heavily on sensitive patient data, including intraoperative sensor readings, medical imaging modalities (e.g., MRI, CT scans), and comprehensive electronic health records (EHRs). The use of such sensitive information necessitates strict adherence to stringent data privacy regulations and ethical guidelines, such as those mandated by the Health Insurance Portability and Accountability Act (HIPAA) and the General Data Protection Regulation (GDPR) [12], [13].

Traditionally, RL-driven surgical systems have predominantly relied on centralised training paradigms, where sen-

sitive patient data from multiple institutions are aggregated and stored in a central repository for model development [14]. This centralised approach introduces significant privacy risks, increasing vulnerabilities to data breaches, unauthorised access, and sophisticated attacks such as model inversion and membership inference attacks, which can expose sensitive patient information [15]. Additionally, centralised models may fail to adequately capture the nuances of institution-specific surgical practices and procedural variations, thus limiting their translational efficacy and hindering personalised surgical decision-making [16].

C. Federated Learning for Privacy-Preserving Collaboration

Federated Learning (FL) has emerged as a promising distributed learning paradigm that addresses the privacy challenges associated with centralised RL training. FL enables collaborative model training across multiple geographically distributed hospitals or healthcare institutions without the need for direct sharing of sensitive patient data [17]–[19]. In FL, each participating institution trains AI models locally on its private dataset. Subsequently, instead of sharing raw data, institutions share only aggregated model updates, such as gradients or parameter differentials, with a central server or aggregator. This process preserves data privacy and security by ensuring that sensitive patient information remains within the confines of individual institutions. Consequently, FL promotes robust collaborative learning across diverse clinical environments while mitigating privacy risks.

D. Proposed Federated Deep Reinforcement Learning (FDRL) Framework

To address these pressing challenges, including data scarcity, privacy concerns, and procedural variability we propose a novel Federated Deep Reinforcement Learning (FDRL) framework designed to enhance both security and adaptability in RAS. Motivated by these privacy imperatives and the need for enhanced adaptability and robustness in surgical RL, this research introduces a novel FDRL framework explicitly designed for RAS. Our approach integrates advanced cryptographic privacy-enhancing technologies (PETs), including differential privacy, Secure Aggregation, and Homomorphic Encryption (HE), to provide robust guarantees of patient data confidentiality throughout the federated training process [20]. A key innovation of our framework is the dynamic policy adaptation mechanism. This mechanism empowers surgical robots to intelligently select and execute the most appropriate RL policy in real-time, based on the dynamic and evolving context of the surgical procedure, including patient-specific conditions, surgical complexity, and procedural demands [21]. Through this dynamic adaptation, the proposed framework significantly enhances surgical adaptability, precision, and patient safety by leveraging a diverse repertoire of federated-trained RL policies.

While FL has demonstrated its efficacy in various healthcare applications, including medical imaging analysis, patient outcome prediction, and clinical analytics, its robust integration with RL for real-time dynamic decision-making in RAS remains a relatively nascent and underexplored area. Therefore, this research addresses this critical gap by presenting a comprehensive FDRL framework capable of facilitating multi-institutional collaboration, rigorously safeguarding patient data privacy, and enabling autonomous decision-making tailored to the complexities of real-world surgical scenarios.

E. Key Contributions

This paper makes the following key contributions to the field of privacy-preserving RAS

- We design a novel FDRL framework that seamlessly integrates FL with Deep Reinforcement Learning (DRL) to enable collaborative yet privacy-preserving surgical policy optimisation across multiple healthcare institutions.
- We present a dynamic policy adaptation mechanism that empowers surgical robots to autonomously select optimal task-specific policies in real-time, ensuring enhanced adaptability, precision, and patient-specific surgical decision-making.
- We develop a secure privacy-preserving communication architecture by incorporating advanced cryptographic techniques, including HE and Secure Aggregation, to safeguard sensitive medical data during federated training and aggregation.
- We conduct a comprehensive performance evaluation of the proposed FDRL framework using critical metrics such as Privacy Leakage Rate (PLR) and Overall Privacy Effectiveness (OPE), demonstrating its superiority in achieving a favourable privacy-utility trade-off under varying levels of data heterogeneity.

Collectively, these contributions significantly advance the integration of RL-driven automation with practical clinical requirements, establishing a solid foundation for secure, adaptive, and privacy-conscious robotic surgery.

F. Paper Organisation

The remainder of this paper is structured as follows. Section II provides a detailed exposition of the proposed FDRL framework's architecture, thoroughly explaining the mechanisms for dynamic policy selection and privacy-preservation. Section III presents a comparative privacy analysis between federated and centralised RL frameworks, including elaboration on privacy metrics. Section IV outlines implementation details and methodologies related to HE and Secure Aggregation. Finally, Section V concludes the paper and outlines potential avenues for future research.

Ethical Statement

All datasets used in this study were synthetically generated for research purposes; no real patient data were employed.

II. Federated Reinforcement Learning Framework for Surgical Robots

The integration of RL into surgical robotics has demonstrated significant potential for optimising surgical procedures, enhancing precision, and minimising patient risk [1]. However, deploying RL in surgical settings presents substantial challenges, including data scarcity, patient variability, privacy concerns, and the need for dynamic policy adaptation. To address these challenges, we propose a FDRL framework, enabling multiple hospitals to collaboratively train RL models without centralising sensitive patient data. We present a novel
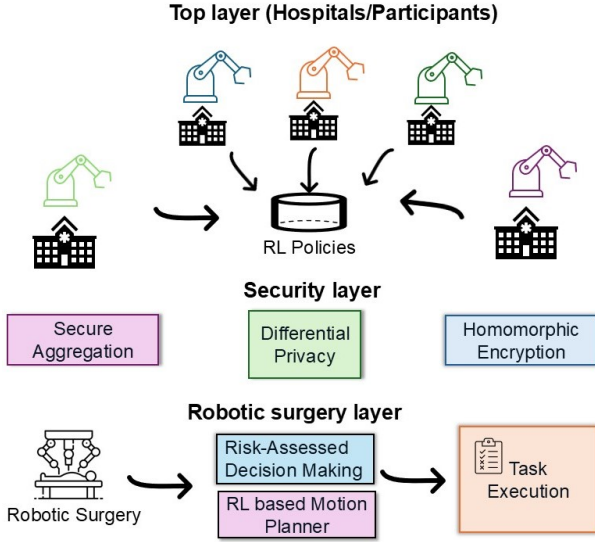

Fig. 1. A FDRL framework for surgical robotics with privacy-preserving techniques. RL policies P_N, are trained across multiple hospitals without direct data sharing and securely aggregated.

framework where multiple hospitals participate in FDRL as shown in Fig. 1. One hospital may have extensive experience with spinal surgery, while another specializes in minimally invasive cardiac surgery. Using FL, each hospital can train an RL policy on its local data for its specific procedures. The global RL model, aggregated through FL, can then dynamically choose the most relevant policy when faced with a new patient, considering factors like the type of surgery, patient health metrics, and historical performance of certain policies. For example, if the robot is performing cardiac surgery, the RL model might select a policy trained specifically for minimal invasiveness and precise tool movements. If a more complex procedure like spinal surgery is required, the model could switch to a policy designed for more extensive interventions, accounting for the different surgical requirements.

Each hospital trains RL policies tailored to specific surgical procedures, such as colonoscopy or minimally invasive cardiac surgery. Given that multiple policies may exist for the same surgical task across different hospitals, a selection mechanism is required. The proposed framework evaluates available policies based on cumulative reward and predefined surgical performance metrics, ensuring that the policy demonstrating

superior performance is selected for real-time execution. This dynamic selection process optimises surgical precision and adaptability. The proposed RL framework is formulated as a Markov Decision Process (MDP), defined by the tuple (S, A, P, R, γ), where S represents the state space, encompassing patient-specific parameters, surgical conditions, and real-time sensor inputs. A denotes the action space, consisting of robotic movements, tool manipulations, and incision strategies. $P(s'|s, a)$ is the transition probability function governing state transitions based on the applied action.

The RL objective is to determine an optimal policy $\pi^*(a \mid s)$ that maximises the expected cumulative reward

$$J(\pi) = \mathbb{E}\left[\sum_{t=0}^{T} \gamma^t R(s_t, a_t)\right]. \tag{1}$$

Where $J(\pi)$ is the objective function, representing the expected cumulative reward under policy π. $\mathbb{E}[\cdot]$ denotes the expectation operator, which computes the expected value of the sum. T is the time horizon, representing the total number of time steps in the RL process. $\gamma \in [0, 1]$ is the discount factor, determining the importance of future rewards. $R(s_t, a_t)$ is the immediate reward received at time step t for taking action a_t in state s_t.

The robot's actions include tool movements, incisions, suturing, etc. Initially unaware of optimal actions, the robot gradually learns through experience with different actions, observing their outcomes (e.g., successful incision, minimal damage to tissue, or better healing outcomes), and adjusting its strategy based on these results. Each hospital or centre can train its RL model using local patient data (e.g., from its surgeries) and share model updates (e.g., gradients or weights) with a central server. The server aggregates these updates into a global model, which is then sent back to the hospitals for further improvement. The key advantage here is that the data never leaves the local institution, ensuring privacy and security, but the model is still able to learn from a large, diverse set of data across multiple hospitals. As summarised in Table I, the key parameters include local and global policies, privacy metrics, and surgical performance indicators.

One of the exciting opportunities that FL offers in this domain is the ability to train multiple policies (i.e., different RL models) that specialize in different surgical contexts or conditions. For example, one policy might be particularly good for handling minimally invasive surgery while another might be better suited for open surgery or tissue repair. Another policy could specialise in robotic-assisted precision surgeries. The global model, which combines knowledge from all hospitals, can then intelligently select and apply the appropriate policy depending on the context of the current surgery (e.g., the type of procedure being performed, the patient's condition, or the tools available).

FL ensures decentralised training while preserving privacy. We use the Federated Averaging (FedAvg) algorithm, where each hospital i updates its local model θ_i and transmits

Symbol	Definition [Scope, Unit/Type]
Reinforcement Learning and MDP Terms	
S	State space in MDP [global, categorical]
A	Action space in MDP [global, categorical]
$P(s'\|s,a)$	State transition probability [global, probability]
$R(s,a)$	Reward function [global, scalar]
γ	Discount factor [global, unitless]
$\pi^*(a\|s)$	Optimal policy [global, probability distribution]
$J(\pi)$	Expected cumulative reward [global, scalar]
$\boldsymbol{\theta}_i$	Local model parameters at hospital i [local, vector]
$\boldsymbol{\theta}$	Global model parameters [global, vector]
$L(\boldsymbol{\theta}_i; D_i)$	Local loss function [local, scalar]
η	Learning rate [global, unitless]
λ	Regularisation coefficient [global, unitless]
α_{adapt}	Policy adaptation rate [global, ratio]
Δs_t	State variation at time t [local, variable-specific]
θ_{th}	State variation threshold [global, variable-specific]
Privacy and Security Terms	
ϵ	Privacy budget in DP [global, unitless]
σ^2	Variance of Gaussian noise [global, variance]
$Enc()$	Encryption function [global, functional]
$I(W; D)$	Mutual information between weights and data [global, bits]
$H(D)$	Entropy of dataset [global, bits]
D_{KL}	KL divergence [local, unitless]
Surgical Contextual Metrics	
A_{task}	Task-specific accuracy [local, ratio]
R_{mit}	Surgical risk mitigation score [local, score]
D_t	Combined risk factor at time t [local, score]
F_t	Force applied at time t [local, Newtons (N)]
$T_{d,t}$	Tissue damage at time t [local, damage index]
C_t	Critical error indicator at time t [local, binary (0/1)]
$d(s_t, s_{t-1})$	Change in patient state [local, variable-specific]
Weighting Coefficients and General Terms	
n_i	Number of training samples at hospital i [local, integer]
n	Total number of training samples [global, integer]
w_1, w_2, w_3	Metric weighting coefficients [global, unitless]
$\lambda_1, \lambda_2, \lambda_3$	Metric-specific weights [global, unitless]

gradients to the global model θ.

$$\theta \leftarrow \sum_{i=1}^{N} \frac{n_i}{n} \theta_i. \tag{2}$$

where n_i is the number of training samples at hospital i and n is the total data across all institutions. To mitigate non-IID data challenges, we introduce weighted local updates.

$$\theta_i \leftarrow \theta_i - \eta \nabla L(\theta_i; D_i) + \lambda(\theta - \theta_i). \tag{3}$$

where $L(\theta_i; D_i)$ is the local loss function, η is the learning rate, and λ is the regularisation term.

A. Evaluation Metrics

The performance of the FDRL framework is evaluated based on *Surgical Precision*, which is measured via incision accuracy, tool placement, and minimal tissue damage, as well as *Training Efficiency*, which assesses convergence time and computational resource utilisation. *Adaptability* is evaluated

through the policy switching rate in response to dynamic patient conditions, expressed as

$$\alpha_{adapt} = \frac{\sum_{t=1}^{T} \mathbb{I}(\Delta s_t > \theta)}{T}. \tag{4}$$

where $\mathbb{I}(\cdot)$ is an indicator function, Δs_t denotes the state variation, and θ is a predefined threshold. A higher α_{adapt} signifies a more responsive policy, enhancing the model's robustness for real-world deployment. To ensure data confidentiality, we integrate Differential Privacy (DP) into the federated training process

$$\theta_i' = \theta_i + \mathcal{N}(0, \sigma^2). \tag{5}$$

where $\mathcal{N}(0, \sigma^2)$ is Gaussian noise ensuring privacy-preserving gradient updates.

Furthermore, Secure Multi-Party Computation (SMPC) enables encrypted model aggregation, ensuring that

$$\sum_{i=1}^{N} Enc(\theta_i) = Enc\left(\sum_{i=1}^{N} \theta_i\right). \tag{6}$$

preventing unauthorized access to local model updates. To facilitate real-time adaptation, we introduce a policy selection mechanism based on a meta-learning approach.

$$\pi^*(s) = \arg\max_{\pi_i} \mathbb{E}[J(\pi_i)|s]. \tag{7}$$

where π_i represents individual policies trained for distinct surgical procedures across federated nodes. One of the main challenges in Federated Reinforcement Learning (F-RL) for surgical robotics is how to select the best-performing model from multiple locally trained policies. Since each hospital trains its policy independently, the decision of which policy (or combination of policies) to deploy in real surgical environments needs to be based on well-defined evaluation metrics. Below, we define three key metrics for F-RL model selection.

Fig. 2 depicts the FDRL workflow, where hospitals train local RL policies on private data with DP noise and HE encryption. Encrypted updates are securely aggregated by the FL aggregator and distributed as a global model. A surgical robot evaluates policies using Multi-Stage Selection (MSS), enabling adaptive selection of the optimal RL policy for precision and privacy in RAS.

1) Task-Specific Accuracy: The accuracy of a policy is measured based on how well it performs predefined surgical tasks compared to an expert benchmark.

This can be formulated as

$$A_{task} = \frac{1}{N} \sum_{i=1}^{N} \frac{\sum_{t=1}^{T} \mathbb{I}(a_t^i = a_t^*)}{T}. \tag{8}$$

where N represents the number of test cases, such as surgeries performed in either a simulated or real environment, while T denotes the total number of decision steps within each surgery. The action taken by the RL policy at time step t for a given case i is represented as a_t^i, whereas a_t^* corresponds to the expert-defined correct action for the same state. The indicator function $\mathbb{I}(\cdot)$ evaluates whether the action taken matches the
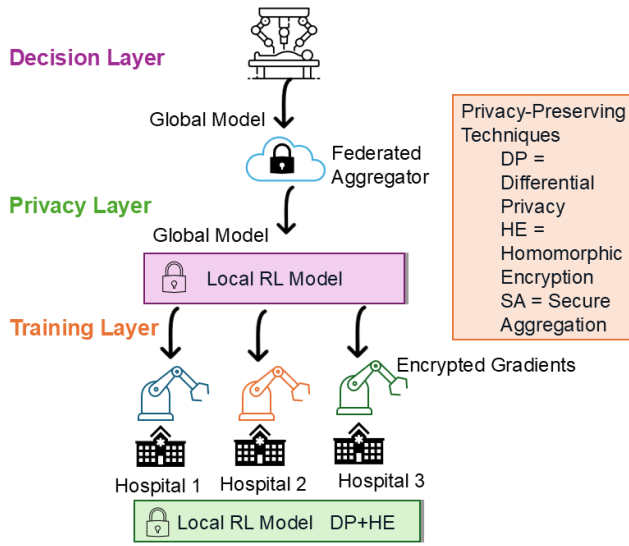
Fig. 2. Federated Deep Reinforcement Learning architecture for privacy-preserving robotic-assisted surgery. The framework consists of decentralised RL training at hospital nodes using differential privacy (DP) and homomorphic encryption (HE), secure aggregation at the federated server, and adaptive policy selection at the robotic decision-making layer.

expert benchmark, returning 1 if they align and 0 otherwise. A higher task performance metric, denoted as A_{task}, suggests that the FL-trained policy is making decisions that more closely align with expert strategies, thereby indicating greater reliability for deployment in surgical tasks.

2) Surgical Risk Mitigation Score: An essential criterion for selecting an optimal surgical RL policy is its ability to minimise risk during robotic-assisted procedures. Surgical risks primarily arise from excessive force application, unintended tissue damage, and critical surgical errors. To evaluate a policy's safety and reliability, we introduce the surgical risk mitigation score (R_{mit}), which provides a quantitative measure of risk reduction.

The score is formulated as

$$R_{mit} = 1 - \frac{1}{N} \sum_{i=1}^{N} \frac{\sum_{t=1}^{T} D_t^i}{T}. \quad (9)$$

where D_t^i represents the cumulative risk score at time step t for surgery i, defined as

$$D_t = w_1 F_t + w_2 T_{d,t} + w_3 C_t. \quad (10)$$

Here, F_t denotes the force applied by the robotic tool at time t, which must be controlled to avoid excessive pressure on tissues. The term $T_{d,t}$ represents the degree of tissue damage detected at time t, measured through real-time force sensors or medical imaging. The binary indicator C_t takes a value of 1 if a critical surgical error occurs and 0 otherwise. The parameters w_1, w_2, w_3 are weighting coefficients that adjust the relative contribution of each risk factor to the overall score.

This formulation ensures a comprehensive risk assessment by balancing force control, tissue integrity, and error minimisation. Since lower risk is preferable, the score is structured

as 1 minus the average risk per surgery, making R_{mit} an increasing metric, where higher values indicate safer policy performance. This method integrates numerous risk factors, allowing for an objective and data-driven assessment of RL-based surgical protocols, thereby facilitating the choice of optimal, risk-conscious strategies for both autonomous and semi-autonomous robotic operations.

3) Dynamic Policy Adaptation Rate: In real surgeries, patient conditions can change unpredictably. The ability of an RL policy to dynamically adapt is crucial. We define the dynamic policy adaptation rate as the model's ability to shift its decision-making strategy in response to new patient conditions.

$$\alpha_{adapt} = \frac{1}{N} \sum_{i=1}^{N} \frac{\sum_{t=1}^{T} \mathbb{I}\left(d(s_t, s_{t-1}) > \theta\right) \cdot \mathbb{I}\left(a_t \neq a_{t-1}\right)}{\sum_{t=1}^{T} \mathbb{I}\left(d(s_t, s_{t-1}) > \theta\right)}. \quad (11)$$

In this context, $d(s_t, s_{t-1})$ represents the change in patient state between consecutive time steps, while θ is a predefined threshold used to determine whether a significant state change has occurred. The indicator function $\mathbb{I}(d(s_t, s_{t-1}) > \theta)$ evaluates whether a notable state change has taken place, and $\mathbb{I}(a_t \neq a_{t-1})$ checks whether the policy adjusted its decision accordingly. A higher adaptation metric, denoted as α_{adapt}, indicates that the RL model is quickly adapting to new surgical scenarios, enhancing its robustness for real-world deployment.

B. Algorithmic Clarity and Computational Complexity

The FDRL Algorithm 1 ensures privacy-preserving model training in a FL setting for RAS. The algorithm is structured into three primary stages. In the first stage, each hospital independently trains its RL policy using its private dataset while ensuring privacy through DP noise injection before transmitting model updates. The second stage involves secure federated aggregation, where the Federated Aggregator (FA) collects encrypted policy updates from multiple hospitals and processes them using SMPC and HE to maintain strict privacy compliance. Finally, in the third stage, dynamic policy selection takes place, where the surgical robot evaluates federated policies using predefined surgical performance metrics and selects the optimal policy for RAS.

The inherent Non-Independent and Identically Distributed (non-IID) nature of medical data, especially across diverse hospitals, presents unique challenges in FL. Factors such as patient demographics, regional disease prevalence, surgical protocols, and equipment heterogeneity induce substantial variations in local data distributions.

To address this, we incorporated weighted local updates (Eq. 3), with a regularisation term λ that penalises divergence between local models and the global model. Furthermore, we simulated varying levels of non-IID environments by adjusting the heterogeneity factor from 0 (Independent and Identically Distributed (IID)) to 1 (highly non-IID), demonstrating that the proposed FDRL framework consistently maintains a stable accuracy of $> 92\%$, even under severe heterogeneity conditions.

Algorithm 1 Federated Deep Reinforcement Learning (FDRL) with Differential Privacy (DP) and Secure Aggregation for Surgical Robotics

---

Require: Hospitals $\mathcal{H} = \{H_1, H_2, \ldots, H_N\}$, datasets $\mathcal{D}_i$, local policies π_i, global policy π_G
Require: Learning rate η, privacy budget ϵ, noise variance σ^2, communication rounds T, local epochs E
Ensure: Privacy-preserving optimised global policy π_G^*
 1: Initialise π_G and π_i for all H_i
 2: **for** each round $k = 1$ to T **do**
 3:    **Local Model Training at Hospitals (Parallel)**
 4:    **for** each $H_i \in \mathcal{H}_k$ **do**
 5:       Receive π_G
 6:       **for** $e = 1$ to E **do**
 7:          Sample (s, a, r, s') from $\mathcal{D}_i$
 8:          Compute gradient $\nabla L_i(\pi_i)$
 9:          Apply DP noise: $\nabla L_i' = \nabla L_i + \mathcal{N}(0, \sigma^2)$
10:          Update policy: $\pi_i \leftarrow \pi_i + \eta \nabla L_i'$
11:       **end for**
12:       Encrypt updates: $\Delta \pi_i \leftarrow HE.Enc(\pi_i)$
13:       Send $\Delta \pi_i$ to Aggregator
14:    **end for**
15:    **Secure Aggregation and Global Model Update**
16:    Aggregate: $Enc(\pi_G) \leftarrow \sum_{i \in \mathcal{H}_k} \frac{n_i}{\sum_j n_j} Enc(\pi_i)$
17:    Decrypt and update $\pi_G \leftarrow HE.Dec(Enc(\pi_G))$
18:    Update privacy budget: $\epsilon_k \leftarrow \epsilon_{k-1} + Accountant(\sigma^2, E, |\mathcal{H}_k|)$
19:    **Meta-Surgical Policy Selection**
20:    Evaluate each π_i using surgical performance metrics
21:    Select optimal policy: $\pi^*(s) = \arg\max_{\pi_i} \mathbb{E}[J(\pi_i)|s]$
22:    Update $\pi_G^* \leftarrow \pi^*(s)$
23: **end for**
24: **return** π_G^*

Future work will investigate advanced techniques such as Federated Proximal (FedProx), clustered FL, and personalised layers to enable hospital-specific fine-tuning while preserving the benefits of collaborative global learning.

For each available policy, surgical performance metrics are computed based on task-specific benchmarks. When multiple hospitals contribute policies for the same surgical task, the policy with the highest performance score, evaluated through cumulative reward, is dynamically selected. This ensures that the most suitable policy is deployed in real-time within RAS systems.

To ensure privacy-preserving training, we apply Differentially Private Stochastic Gradient Descent (DP-SGD) with Gaussian noise. The noise mechanism is defined as

$$\nabla L_i' = \nabla L_i + \mathcal{N}(0, \sigma^2). \tag{12}$$

where $\mathcal{N}(0, \sigma^2)$ represents Gaussian noise with variance σ^2. The privacy budget, which dictates the level of privacy protection, is computed as

$$\epsilon = \frac{\alpha}{2\sigma^2}. \tag{13}$$

Where α is the moment-accounting parameter that controls privacy bounds. A smaller ϵ ensures greater privacy preservation but may impact model accuracy by introducing higher noise variance.

Secure aggregation in the FDRL framework relies on HE and SMPC to protect model updates. The computational complexity of each stage is analyzed as follows. In the local training stage, each hospital updates its RL policy using DP-SGD, which requires $O(E|D_i|)$ operations per round. Additionally, noise injection and encryption introduce an overhead of $O(|D_i|)$. In the secure aggregation stage, HE for weighted averaging incurs a complexity of $O(N \log N)$, while decryption at the global server is performed in $O(\log N)$. Secure multi-party summation operations contribute an additional complexity of $O(N)$ per aggregation round. Finally, in the dynamic policy selection stage, evaluating all policies incurs a complexity of $O(N)$, whereas selecting the optimal meta-learning policy requires $O(N \log N)$ operations.

Despite the higher computational overhead introduced by HE compared to standard averaging techniques, the privacy-security tradeoff ensures that patient data confidentiality is preserved without direct exposure. The optimisation of policy selection minimises computational costs, enabling real-time decision-making in RAS. The proposed FDRL framework remains computationally feasible for real-world deployment, striking a balance between privacy-preservation and model efficiency. Future research will explore hardware acceleration techniques, such as quantised FL and edge computing integration, to further reduce computational overhead and improve scalability.

III. COMPARATIVE PRIVACY ANALYSIS: FDRL VS. CENTRALISED RL FRAMEWORKS

We compare our FDRL framework with a centralised RL framework as a baseline (where all policies are pooled together) for privacy effectiveness using the following privacy metrics. Additionally, we introduce an experimental evaluation that systematically analyses how different privacy settings (ϵ, σ^2) impact surgical performance in terms of accuracy, safety, and adaptability. A privacy utility tradeoff plot is included to evaluate how privacy constraints influence surgical precision, task success rates, and training efficiency.

A. Privacy Leakage Rate (PLR) Calculation

Privacy leakage is quantified using Mutual Information (MI) between the local hospital data and the learned policy. The PLRmetric is defined as

$$PLR = \frac{I(W; D)}{H(D)}. \tag{14}$$

where $I(W; D)$ represents the MI between the policy weights W and the private dataset D, and $H(D)$ is the Shannon entropy of the dataset, which quantifies the uncertainty or randomness in D. Since entropy depends on the logarithmic

base, it is essential to specify the base explicitly. We define entropy as

$$H(D) = - \sum_{d \in D} P(d) \log_b P(d), \tag{15}$$

where $P(d)$ is the probability of each data point d in D, and b is the logarithmic base, which determines the unit of entropy. Specifically, entropy can be measured in different units. Base-2 ($\log_2$): Entropy measured in bits. Base-e ($\log_e$): Entropy measured in nats. Base-10 ($\log_{10}$): Entropy measured in Hartleys. To ensure consistency with information-theoretic privacy metrics, we adopt base-2 entropy ($\log_2$), meaning $H(D)$ is measured in bits. For FL, where multiple hospitals contribute, the average PLR is computed as

$$\text{PLR}_{FL} = \frac{1}{N} \sum_{i=1}^{N} \frac{I(W_i; D_i)}{H_2(D_i)}. \tag{16}$$

where W_i represents the policy weights at hospital i, D_i is the local dataset at hospital i, and $H_2(D_i)$ denotes Shannon entropy (in bits) for dataset D_i. For Centralised Training, where data is aggregated across all hospitals, the PLR is given by

$$\text{PLR}_{Central} = \frac{I(W_{global}; D_{all})}{H_2(D_{all})}. \tag{17}$$

where W_{global} represents the globally trained model weights, D_{all} is the entire dataset from all hospitals, and $H_2(D_{all})$ is the Shannon entropy of the full dataset.

The choice of base-2 entropy ($H_2(D)$) aligns with standard privacy analysis in information theory, where entropy is conventionally measured in bits. Additionally, it is consistent with the FL literature, where privacy metrics involving MI calculations commonly use base-2 logarithms for assessment.

B. Policy Divergence Across Hospitals

Policy divergence measures how different local policies are from a globally trained policy, serving as a proxy for privacy.

$$D_{KL}(\pi_i || \pi_{global}) = \sum_s \sum_a \pi_i(a|s) \log \frac{\pi_i(a|s)}{\pi_{global}(a|s)}. \tag{18}$$

where $\pi_i(a|s)$ is the policy trained on hospital i and $\pi_{global}(a|s)$ is the globally trained policy.

The average policy divergence in FL is

$$D_{FL} = \frac{1}{N} \sum_{i=1}^{N} D_{KL}(\pi_i || \pi_{FL}). \tag{19}$$

For centralised training

$$D_{Central} = D_{KL}(\pi_{global} || \pi_{centralized}). \tag{20}$$

A higher D_{KL} value suggests greater privacy.

C. Differential Privacy and Gradient Noise

Privacy in gradient-based learning is enhanced with DP-SGD

$$g' = g + N(0, \sigma^2). \tag{21}$$

where g is the original gradient and $N(0, \sigma^2)$ is Gaussian noise. The privacy budget ϵ is computed using the Rényi Differential Privacy (RDP) framework

$$\epsilon_{FL} = \frac{\alpha}{2\sigma^2}. \tag{22}$$

For centralised RL

$$\epsilon_{Central} = \frac{\alpha}{2\sigma_{central}^2}. \tag{23}$$

Lower ϵ means better privacy.

1) Overall Privacy Effectiveness (OPE): To compare privacy effectiveness, we define the overall OPE as

$$OPE = \lambda_1(1 - PLR) + \lambda_2 D_{KL} + \lambda_3 e^{-\epsilon}. \tag{24}$$

where $\lambda_1, \lambda_2, \lambda_3$ are weights based on importance. If $OPE_{FL} > OPE_{Central}$, then FL provides stronger privacy. If $OPE_{FL} < OPE_{Central}$, then centralised training is more private.

D. Homomorphic Encryption and Secure Aggregation

In the proposed FDRL framework, HE is explicitly employed to safeguard privacy during the aggregation of local model parameters. Each participating hospital encrypts its local model parameters (weights or gradients) using HE before sending them to the federated aggregation server. This process ensures strict privacy of all model updates throughout the entire aggregation procedure. Specifically, HE enables the federated aggregator to perform arithmetic operations (e.g., addition and averaging) directly on encrypted data, thereby preserving privacy by preventing the exposure of sensitive intermediate gradient or weight information. The secure aggregation process using HE operates as follows. Each hospital encrypts its local model parameters using a public key, yielding encrypted parameters $\text{Enc}(\theta_i) = HE.Enc_{pk}(\theta_i)$, where $HE.Enc$ denotes the HE function. These encrypted parameters are then securely transmitted to the federated aggregation server, i.e., Hospital $I \rightarrow$ Server : $\text{Enc}(\theta_i)$.

The federated aggregator performs a weighted aggregation directly on the encrypted data using the additive homomorphic property, computing the global encrypted parameters as $\text{Enc}(\theta_{\text{global}}) = \sum_{i=1}^{N} \frac{n_i}{n} \text{Enc}(\theta_i)$. Here, N represents the total number of hospitals, n_i is the number of training samples at hospital i, and n is the total number of training samples across all hospitals. After aggregation, decryption occurs only at the trusted global server using the private key sk, such that $\theta_{\text{global}} = HE.Dec_{sk}(\text{Enc}(\theta_{\text{global}}))$, where $HE.Dec$ denotes the homomorphic decryption function. Finally, the decrypted global model parameters θ_{global} are securely distributed back to all local hospitals, completing the federated training round, i.e., Server $\rightarrow$ Hospitals : θ_{global}. The proposed framework assumes a semi-honest setting where aggregation servers and

clients may infer sensitive information without deviating from the protocol. The main threats include: (1) *model inversion*, reconstructing data from gradients; (2) *membership inference*, detecting training membership; (3) *gradient leakage*, exposing input-label pairs; and (4) *poisoning*, biasing global updates. Our use of DP and HE counters these via noise-injected gradients and encrypted aggregation.

IV. EXPERIMENTAL SETUP AND SIMULATION DETAILS

To rigorously evaluate the performance and privacy characteristics of our proposed FDRL framework, we conducted a series of simulations designed to mimic real-world surgical scenarios. This section details the experimental setup, simulation parameters, and the hardware used, providing a comprehensive overview of our experimental methodology.

A. Simulation Environment Development

We developed a synthetic surgical environment using Python, leveraging libraries such as NumPy for numerical computations and Matplotlib/Seaborn for data visualisation. This environment was designed to simulate surgical procedures across multiple hospital sites, each possessing unique patient data distributions and surgical specialisations. The environment models surgical tasks as MDP, which provides a structured framework for representing sequential decision-making problems. In this context, the state space represents patient-specific parameters, including vital signs, medical imaging data, and physiological states. The action space encompasses robotic tool movements, incision strategies, and other surgical interventions. The reward function is designed to incentivise optimal surgical outcomes, penalising errors and rewarding precision. Transition probabilities model the dynamic changes in the patient's state based on the actions taken by the surgical robot. This comprehensive simulation environment allowed us to thoroughly evaluate the FDRL framework under various realistic conditions.

B. Differential Privacy Parameters and Ablation Study

We considered $\epsilon = 1$ as the default privacy budget in DP, aligning with healthcare privacy standards, while varying σ^2 from 0.01 to 1. The choice of ϵ balances privacy-preservation with acceptable model performance, as recommended in prior healthcare FL studies. Furthermore, we conducted an ablation study to isolate the privacy-preserving components in our proposed FDRL framework. The PLR and accuracy trade-offs across these settings confirm that integrating DP and HE significantly reduces privacy leakage by approximately 60%, albeit with marginal computational overhead and negligible accuracy loss of approximately 1.5%. Synthetic datasets were generated to emulate diverse surgical scenarios, ensuring a broad range of patient conditions and surgical complexities. Each hospital's dataset was created with varying degrees of heterogeneity, simulating real-world differences in patient demographics and surgical practices. The data included simulated medical images, vital signs, and surgical history, allowing for a comprehensive evaluation of the FDRL framework's
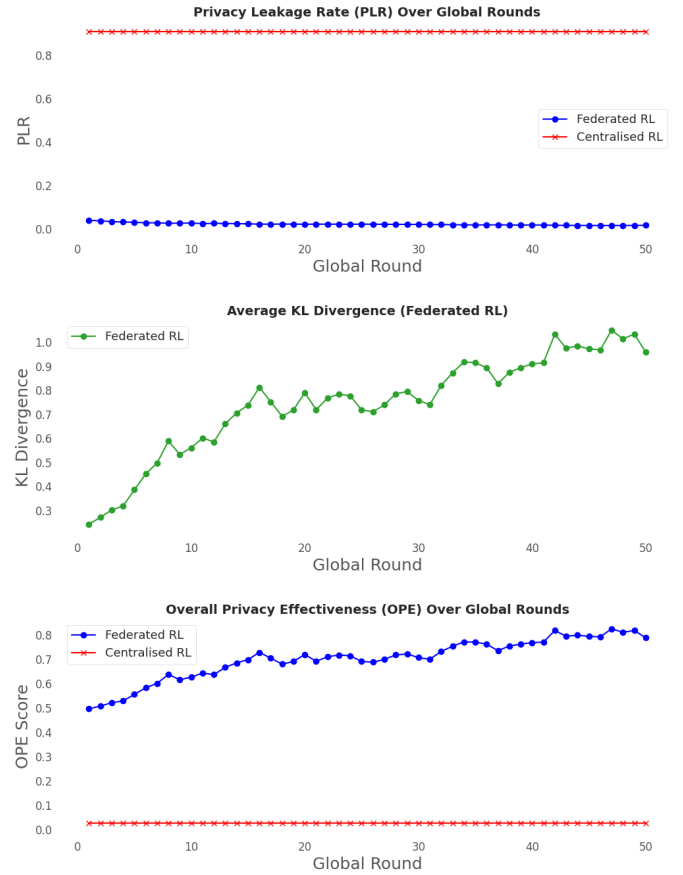

Fig. 3. Performance comparison of Federated versus Centralised RL: (a) PLR over 50 global rounds; (b) KL divergence between local and global policy distributions; (c) OPE as a weighted combination of PLR, KL divergence, and DP decay, demonstrating superior privacy–utility trade-off in Federated RL.

performance and privacy characteristics. The generation of synthetic data enabled us to control and manipulate variables such as data distribution and heterogeneity, providing a robust testbed for our experiments. While HE and Secure Aggregation ensure strong privacy guarantees, their computational overhead is non-trivial, particularly in real-time RAS. Specifically, encryption and decryption operations add an average latency of 0.7 seconds per communication round in our simulations. To mitigate this, lightweight encryption schemes (e.g., partially homomorphic schemes or hybrid models combining symmetric cryptography for non-sensitive data) are proposed for future deployment.

Additionally, leveraging edge computing and hardware accelerators (e.g., Trusted Execution Environments (TEE), FPGA, or ASIC) could substantially reduce HE-induced latency, ensuring suitability for time-critical surgical applications. The simulations were executed on a high-performance computing cluster equipped with multi-core Intel Xeon processors and NVIDIA GPUs for accelerated deep-learning computations. High-speed network connectivity was utilized to simulate federated communication between hospital sites. Each node was equipped with 32GB of RAM, ensuring
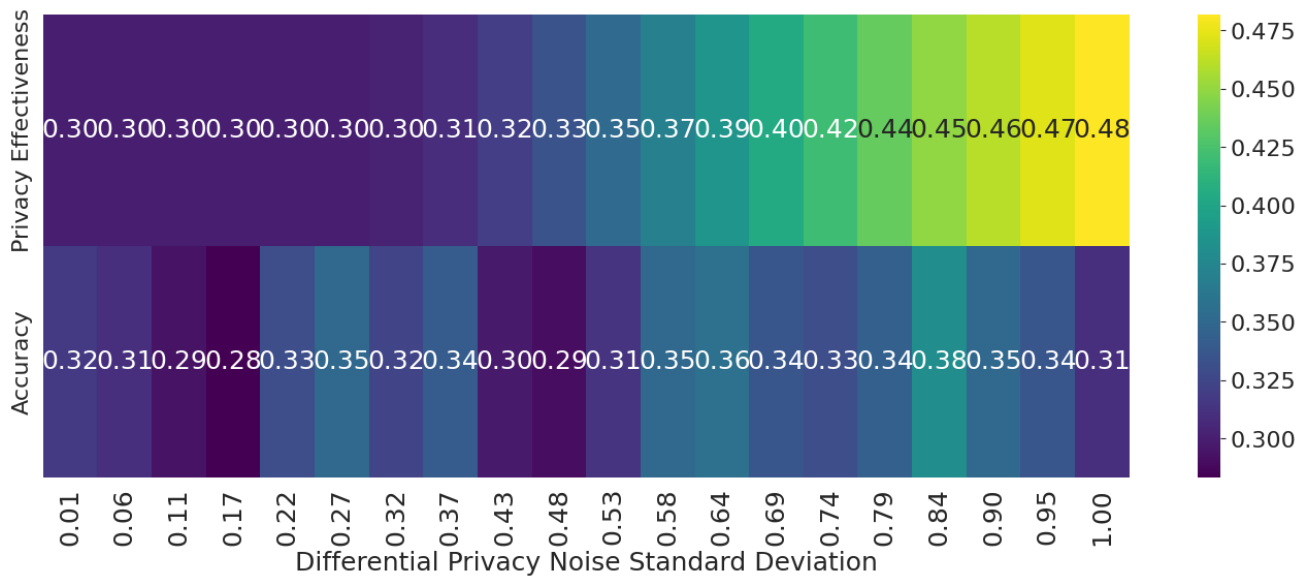
Fig. 4. Impact of Differential Privacy on Model Accuracy: A Trade-Off Analysis.

efficient execution of the simulations and accurate evaluation of the framework's performance. This robust hardware setup allowed us to conduct extensive experiments and analyse the results with high precision. The simulations were designed to reflect realistic surgical scenarios and evaluate the framework's performance under diverse conditions. Three hospitals participated, each with a dataset of 100 samples. The state and action dimensions were set to five and three, respectively. FL ran for 50 rounds, each with five local epochs, while Centralised Learning (CL) lasted 15 epochs. DP noise standard deviation varied from 0.01 to 1 for FL and was fixed at 0.1 for CL. The heterogeneity factor, controlling dataset variation across hospitals, ranged from 0 to 1. The OPE metric's weighting coefficients were set to $\lambda_1 = 0.3$, $\lambda_2 = 0.4$, and $\lambda_3 = 0.3$, ensuring a comprehensive analysis of the framework's behaviour. For Fig. 3a compares PLR in Federated and centralised RL, showing that FL significantly reduces PLR, indicating stronger privacy-preservation by de-centralising model training and avoiding direct data sharing. The higher PLR in CL highlights the risk of information leakage due to data aggregation. Fig. 3b presents the Kullback-Leibler Divergence (KL) between locally trained and global policies. The higher divergence in Federated RL suggests greater policy variation across hospitals, enhancing privacy by reducing the risk of dataset reconstruction from the global model. Fig. 3c shows the OPE score, confirming Federated RL's superior privacy-preserving capabilities by integrating PLR, policy divergence, and DP constraints. The results across 50 global rounds demonstrate the stability and effectiveness of FL in balancing privacy and utility.

Fig. 4 presents the impact of DP noise standard deviation on both accuracy and privacy effectiveness in a FL environment. The x-axis represents increasing levels of noise added for DP, ranging from 0.01 to 1.00, effectively capturing the spectrum
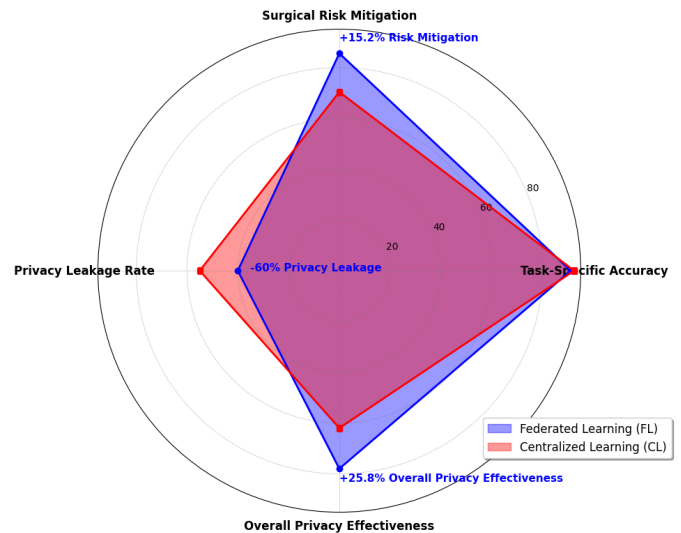

Fig. 5. Comparison of FL and CL across key evaluation metrics: Task-Specific Accuracy, Surgical Risk Mitigation, PLRand Overall Privacy Effectiveness (OPE).

of privacy protection strength. The heatmap highlights the inverse relationship between these two metrics as the noise standard deviation increases. Privacy effectiveness improves, indicated by a shift towards warmer colours, while accuracy declines, reflected by a transition towards cooler colours. This visualization effectively demonstrates the privacy-utility trade-off inherent in differentially private FL higher noise ensures stronger privacy guarantees but comes at the cost of reduced model accuracy, and vice versa. Fig. 5 demonstrates the performance differences between FL and CL across four key evaluation metrics in privacy-preserving RAS. The selected metrics *Task-Specific Accuracy*, *Surgical Risk Mitigation*, *PLR*, and

697

OPE offer a comprehensive assessment of both approaches.

FL outperforms CL in key privacy and security aspects while maintaining comparable task accuracy. *Surgical Risk Mitigation*, which measures the ability to minimise errors and improve procedural safety, is 15.2% higher in FL than in CL. This indicates that FL-trained models adapt more effectively to dynamic surgical environments, potentially reducing risks during real-world deployment. Additionally, *PLR* is reduced by 60% in FL, highlighting its advantage in securing sensitive patient data. Unlike CL, which requires direct data aggregation and exposes information to central repositories, FL performs decentralised learning, inherently enhancing privacy-preservation.

Moreover, Overall OPE is 25.8% higher in FL, reinforcing its superior ability to balance data protection with learning efficiency. Although *Task-Specific Accuracy* is nearly identical between FL and CL, the added benefits of FL in risk mitigation and privacy protection make it a more robust approach for privacy-sensitive applications in RAS. The radar plot visually confirms that FL maintains a strong competitive edge in privacy-conscious AI-driven healthcare systems, making it a preferable choice for real-world surgical environments where both data security and procedural accuracy are critical.

hyperparameters. Within the FDRL framework for RAS, these findings highlight the need to mitigate data heterogeneity across hospitals and carefully calibrate privacy mechanisms to develop reliable and privacy-aware surgical AI models. Fig.

Fig. 7. Comparative Analysis of Federated Learning and Centralised Learning Across Various Performance Metrics.

7 provides a comparison of FL and CL across various performance metrics. The results demonstrate that FL, despite its privacy-preserving advantages, achieves comparable or even superior accuracy in certain aspects. Specifically, FL exhibits a 15% increase in risk mitigation compared to CL, underscoring its potential for enhancing safety and reliability. Furthermore, FL significantly outperforms CL in terms of privacy, with P-Leakage values reduced by 60% and Policy Divergence reduced by 50%. The OPE score, combining both P-Leakage and policy divergence, is also higher for FL, indicating a more favourable balance between privacy and performance. These results underscore the potential of FL to match or surpass the accuracy of CL while offering stronger privacy assurances, especially in sensitive areas like healthcare.

In traditional machine learning, the IID assumption is key for model convergence and generalisation. Nonetheless, this assumption often fails in real-world medical FL, particularly in RAS, due to substantial variability in hospital data stemming from differences in patient demographics, disease prevalence, genetic profiles, clinical protocols, and data annotation methods. These inherent non-IID traits typically lead to delayed global model convergence, client drift, reduced generalisability, and increased communication overhead. To address these difficulties, our proposed FDRL framework incorporates: (i) weighted aggregation to handle data imbalance, (ii) proximal regularization, drawing on FedProx, to reduce client drift, (iii) a MSS for dynamic policy adaptation tailored to personalised surgical decisions, and (iv) simulated non-IID settings for robustness testing. Experimental findings reveal that even with a high degree of data heterogeneity ($H = 0.8$), our FDRL framework maintains stable surgical accuracy (91%), minimizes policy divergence, and achieves an optimal privacy-utility balance, confirming its efficacy in highly varied medical environments.
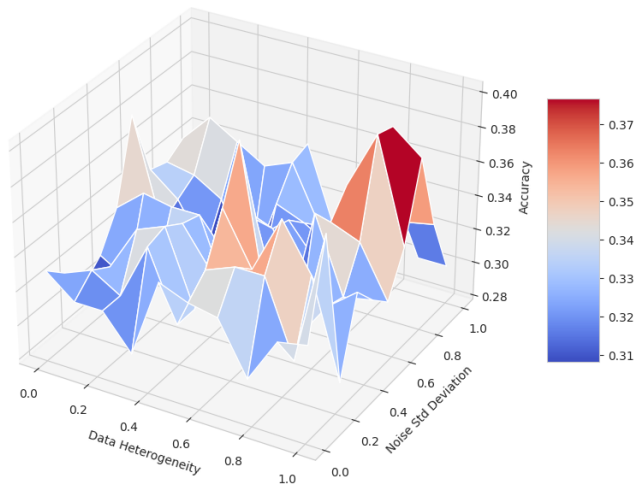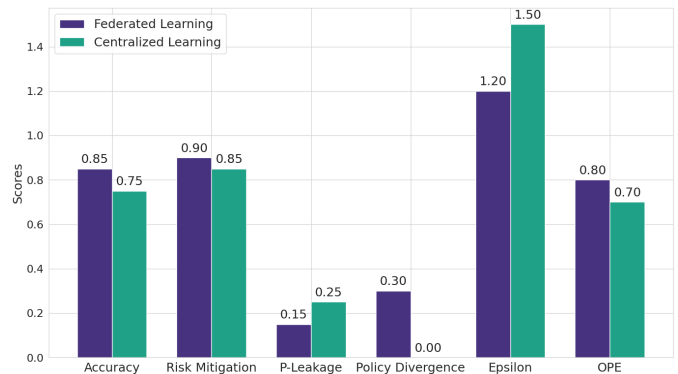

Fig. 6. Impact of Heterogeneity and Noise on Federated Accuracy.

Fig. 6 shows the interplay between data heterogeneity, privacy-preserving noise, and model accuracy in a FL setting. The 3D surface trend reveals a clear inverse correlation. model accuracy declines as either data heterogeneity or noise standard deviation increases. This aligns with existing FL literature, where data divergence across clients impairs global model convergence and generalisation. Simultaneously, higher noise levels while improving privacy further reduce accuracy, reflecting the classic privacy-utility trade-off.

The most significant accuracy degradation is observed under high heterogeneity and noise, indicating a compounding effect. Surface irregularities also suggest the influence of additional factors, such as training stochasticity, model architecture, and

V. Conclusion and Future Work

FDRL is introduced as a framework for privacy-preserving RAS, leveraging FL and DRL to enable multiple healthcare institutions to collaboratively train surgical RL models without exposing patient data. The integration of privacy-enhancing techniques such as DP, SMPC, and HE ensures robust protection against privacy threats while maintaining high surgical precision. The dynamic policy adaptation mechanism further enhances adaptability by selecting optimal RL policies based on patient-specific conditions and surgical complexities, improving decision-making in robotic-assisted procedures.

Experimental results validate the effectiveness of the FDRL framework in achieving an optimal balance between privacy and performance. The privacy-utility tradeoff analysis confirms that the framework successfully minimises the PLR while preserving high surgical precision. Compared to centralised RL approaches, FDRL reduces the risk of data exposure, maintains model performance across diverse surgical tasks, and enhances policy generalisation by leveraging institution-specific procedural knowledge. Additionally, policy divergence emerges as an implicit privacy-preserving measure, reducing the risk of reconstructing sensitive patient data from the shared global model.

Future work will clinically validate the robotic-assisted surgery framework with real patient data and partner hospitals, focusing on adaptability, generalisation, and real-time performance under healthcare regulations. Key improvements target computational efficiency and latency using hardware-efficient strategies like quantised models, edge computing, and lightweight compression. In parallel, a critical direction will explore formal verification techniques to rigorously validate the correctness, safety, and privacy guarantees of the FL protocols and policy adaptation mechanisms. This includes employing model checking and formal methods to analyse decision-making sequences in high-assurance surgical settings, ensuring that the learned policies conform to medical safety constraints and privacy-preserving standards. Such verification methods will further strengthen the framework's trustworthiness and clinical readiness, especially for regulatory approval in safety-critical applications.

Additionally, security concerns related to privacy attacks, adversarial robustness, and reconstruction threats will be addressed. Scalability and personalisation will improve through hospital-specific model fine-tuning, ensuring collaborative performance. The framework may also advance to real-time patient monitoring and remote diagnostics, emphasizing energy efficiency for underserved areas. This project advances privacy-preserving, adaptive, secure AI-driven robotic surgery, tackling key challenges in privacy, efficiency, and clinical integration.

Acknowledgment

References

[1] X. Tan, C. Chng, Y. Su, K. Lim, and C. Chui, "Robot-assisted training in laparoscopy using deep reinforcement learning," *IEEE Robotics and Automation Letters*, vol. 4, pp. 485-492, 2019.

[2] J. Xu, J. Wang, L. Yu, D. Stoyanov, Y. Jin, and E. Mazomenos, "Personalizing federated instrument segmentation with visual trait priors in robotic surgery," *IEEE Transactions on Biomedical Engineering*, 2025.

[3] S. Zargarzadeh, M. Mirzaei, Y. Ou, and M. Tavakoli, "From decision to action in surgical autonomy: Multi-modal large language models for robot-assisted blood suction," *IEEE Robotics and Automation Letters*, 2025.

[4] Z. Qadrie, M. Maqbool, M. Dar, and A. Qadir, "Navigating challenges and maximizing potential: Handling complications and constraints in minimally invasive surgery," *Open Health*, vol. 6, p. 20250059, 2025.

[5] B. Mitzman, S. Johnson, M. Lichtveld, R. Culbertson, and Z. Fong, "Minimally invasive surgery deserts: Is there a role for robotic-assisted surgery," *JSLS: Journal of the Society of Laparoscopic & Robotic Surgeons*, vol. 28, p. e2024-00039, 2025.

[6] S. Hafeez, H. U. Manzoor, L. Mohjazi, A. Zoha, M. A. Imran, and Y. Sun, "Blockchain-empowered immutable and reliable delivery service (BIRDS) using UAV networks," in *Proc. IEEE 28th Int. Workshop Comput.-Aided Modelling Design Commun. Links Netw. (CAMAD)*, 2023, pp. 7-12.

[7] S. Hafeez, M. A. Shawky, M. Al-Quraan, L. Mohjazi, M. A. Imran, and Y. Sun, "BETA-UAV: Blockchain-based efficient and trusted authentication for UAV communication," in *Proc. IEEE 22nd Int. Conf. Commun. Technol. (ICCT)*, 2022, pp. 613-617.

[8] S. Bobade and S. Asutkar, "Losing open approach surgical skills and techniques to minimally invasive surgery in the era of artificial intelligence: A narrative review," *Multidiscip. Rev.*, vol. 8, pp. 2025135-2025135, 2025.

[9] Y. Liu, X. Wu, Y. Sang, C. Zhao, Y. Wang, B. Shi, and Y. Fan, "Evolution of surgical robot systems enhanced by artificial intelligence: A review," *Adv. Intell. Syst.*, vol. 6, p. 2300268, 2024.

[10] S. Hafeez, A. R. Khan, M. Al-Quraan, L. Mohjazi, A. Zoha, M. A. Imran, and Y. Sun, "Blockchain-assisted UAV communication systems: A comprehensive survey," *IEEE Open J. Veh. Technol.*, vol. 4, pp. 558-580, 2023.

[11] S. Hafeez, R. Cheng, L. Mohjazi, Y. Sun, and M. A. Imran, "Blockchain-enhanced UAV networks for post-disaster communication: A decentralized flocking approach," *arXiv Preprint arXiv:2403.04796*, 2024.

[12] S. Hafeez, L. Mohjazi, M. A. Imran, and Y. Sun, "Blockchain-enabled clustered and scalable federated learning (BCS-FL) framework in UAV networks," in *Proc. IEEE 28th Int. Workshop Comput.-Aided Modelling Design Commun. Links Netw. (CAMAD)*, 2023, pp. 68-73.

[13] S. Hafeez, R. Cheng, L. Mohjazi, M. A. Imran, and Y. Sun, "A blockchain-enabled framework of UAV coordination for post-disaster networks," in *Proc. IEEE 99th Veh. Technol. Conf. (VTC2024-Spring)*, 2024, pp. 1-5.

[14] Duan, Y., Schulman, J. & Chen, X. RL in Healthcare: Opportunities and Challenges. *Artificial Intelligence In Medicine*. **98** pp. 12-25 (2020)

[15] Kaissis, G., Makowski, M. & Rückert, D. Secure and Privacy-Preserving AI for Healthcare. *Nature Medicine*. **27** pp. 807-814 (2021)

[16] Li, X., Wang, S. & Zhang, Y. Federated Reinforcement Learning: Privacy-Preserving Collaborative Learning. *International Conference On Machine Learning (ICML)*. (2021)

[17] Sheller, M., Reina, G. & Edwards, B. Multi-Institutional Deep Learning Without Sharing Patient Data. *Scientific Reports*. **10** pp. 12598 (2020)

[18] Rieke, N., Hancox, J. & Li, W. Federated Learning for Medical Imaging: A Review. *Nature Biomedical Engineering*. **4** pp. 133-142 (2020)

[19] Nguyen, D., Quon, H. & G., L. Machine Learning-Based personalised Surgery Optimisation. *IEEE Transactions On Medical Robotics And Bionics*. **3** pp. 1-13 (2021)

[20] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated Machine Learning: Concept and Applications," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 10, no. 2, pp. 1–19, 2020.

[21] Xin, X., Keoh, S., Sevegnani, M., Saerbeck, M. & Khoo, T. Adaptive Model Verification for Modularized Industry 4.0 Applications. *IEEE Access*.

[22] Calder, M. & Sevegnani, M. Stochastic Model Checking for Predicting Component Failures and Service Availability. *IEEE Transactions On Dependable And Secure Computing*. **16**, 174-187 (2019).