Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

Robust Tumor-Immune System Regulation With Drug Administration Via Reinforcement Learning

V Abhilash Reddy[1], G Shruthik[2], B Akhil[3]
[123]Department of IT, SNIT, Telangana, India

Abstract **- In cancer immunotherapy, it is important to have ideal control of the dosage of drugs used to prevent the proliferation of cancer without affecting the stability of the immune systems and preventing treatment toxicity. Conventional control-based methods are based on correct mathematical models and fixed parameters, which in many cases are hardly accessible as a result of nonlinear and uncertain tumour–immune interactions. In this study, we propose a deep reinforcement learning (DRL)-based adaptive robust drug administration control (DRL-ARDAC) that can be used as an intelligent solution to drug administration challenges in cancer management. The proposed system describes tumour–immune interactions with a nonlinear augmented state representation that incorporates tumour and immune cells and tracking error. An adaptive policy of drug dosage is determined with the help of a reinforcement learning-based critic neural network that approximates the optimal value function. Moreover, a non-quadratic costing function with a discounted cost is proposed to impose limited dose constraints of drugs and to control the toxicity of the treatment. Simulation experiments demonstrate that the proposed DRL-ARDAC framework provides better tumour suppression (95%), less drug use (0.58 units), improved immune stability (93%), and reduced computation time (1.5s) in comparison with traditional control techniques. These findings reveal that the developed strategy is effective and reliable for smart and customized cancer immunotherapy.**

Keywords - Reinforcement Learning, Drug Dosage Control, Tumor–Immune System Modeling, Immunotherapy Optimization, Neural Network Control

I. INTRODUCTION

Cancer is among the major causes of death in every corner of the globe, and the biological aspects associated with cancer are so complex that it becomes extremely hard to treat. The commonly used traditional methods of treating cancer, like chemotherapy, radiotherapy, and surgery, are mainly intended to destroy the tumour cells. Still, in most cases, the treatment has serious side effects because of the destruction of normal tissues. Cancer immunotherapy has been a promising treatment approach that has developed over the last few years, improving the body's immune system to identify and eliminate cancerous cells. Unlike standard treatments, immunotherapy aims to boost immune responses to inhibit tumour growth without disrupting the balance between immune cells in the body [1]. Nevertheless, the dosing of drugs and the dynamics of the immune response should be carefully controlled to achieve optimal therapeutic results. The interaction between immune cells and tumour cells is highly nonlinear and is determined by several biological unknowns. Tumour-immune dynamics have been extensively analysed through mathematical modelling and the design of treatment strategies. The Kirschner-Panetta model is an early tumour-immune model that identifies the interactions among tumour cells, effector immune cells, and cytokines using nonlinear differential equations [2]. The models offer a theoretical explanation of cancer development and an assessment of therapeutic interventions. Similarly, several studies have used optimal control theory to determine the most effective policies for drug administration to minimise tumour growth and drug toxicity [3]. Nonetheless, traditional control methods require very precise system parameters and very simplistic assumptions, limiting their use in practice in a clinical environment. The second issue in immunotherapy treatment is the limitation of drug dosage. Overdose can cause toxicity, immunosuppression or serious side effects, and a lower dosage does not ensure that the tumour is managed. Most classical control methods are either insensitive to dosage limits or unable to adapt to patient-specific differences. Since cancer progression and immune response vary across patients, the use of individualised treatment plans is necessary to achieve safe and effective treatment [4]. As a result, scientists have begun examining intelligent, data-driven technologies that can adapt to the selection of treatment options in dynamic, uncertain biological settings.

Over the past few years, reinforcement learning (RL) has attracted significant interest for solving complex control problems in which the system dynamics are not fully known and are uncertain. The use of RL enables an intelligent agent to apply optimal decision-making policies by interacting with the environment, maximising cumulative rewards [5]. This has been further augmented by breakthroughs in deep reinforcement learning (DRL) which combines neural networks with reinforcement learning algorithms to enable the system to estimate complex value functions and control policies [6]. DRL has been effectively implemented in other fields of application, including robotics, autonomous systems, and health care decision support [7]. When applied to cancer therapy, reinforcement learning can learn the optimal drug administration strategy on its own and consistently adapt to changes in tumour growth and the immune response. The use of RL in healthcare treatment planning, such as personalised drug dosing and dynamic treatment regimes, has been the subject of several recent studies. These methods indicate that RL can enhance clinical decision-making to optimise treatment options based on patients' specific characteristics and system feedback [8]. In addition, the control methods of adaptive dynamic programming and RL have been

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

demonstrated to offer robust solutions for nonlinear systems with uncertainties and disturbances [9]. Nevertheless, current RL-based cancer treatment methods either make optimistic assumptions about the system or focus on regulatory issues rather than trajectory tracking and limitations in drug dosage.

To overcome these shortcomings, this work presents a DRL-ARDAC framework for tumour-immune system control. The system describes the dynamics between tumour and immune cells using a nonlinear augmented state representation and a critic network that employs reinforcement learning to estimate optimal control policies. The proposed framework incorporates the limited drug dosage restrictions into a discounted cost, which guarantees safe and effective treatment plans and a reduction in drug toxicity. The adaptive learning system also helps modify treatment policies in response to real-time conditions, including biological uncertainty and perturbations. The proposed solution is expected to ensure effective tumour suppression without compromising immune system stability, based on intelligent decision-making and robust control, thereby facilitating safer and personalised control of cancer immunotherapy.

Section II of this paper reviews related work, Section III presents the problem, Section IV presents the proposed methodology, Section V presents the experiments and results, and finally, the conclusion is found in Section VI, including future directions.

II. RELATED WORKS

In [11], the authors came up with a new goal-directed reinforcement learning (GORL) system to optimise the control of drug dosages in critically ill patients, especially those affected by sepsis. The process of making decisions regarding drug dosage in the intensive care unit is a complicated sequential process in which clinicians have to vary the amount of medication according to the patient's conditions constantly. Rule-based or more traditional optimisation techniques tend to be difficult to use in cases involving delayed rewards, large states, and changing clinical conditions. To address these problems, the authors propose a hierarchical reinforcement learning framework that decomposes the overall treatment goal into several short-term objectives. The proposed GORL model is a two-agent cooperative learning model that simulates clinicians' decision-making and provides a dosage strategy to achieve long-term patient recovery. A goal-directed mechanism and an intrinsic, hindsight-rewarding mechanism are also added to enhance learning efficacy and address the issue of sparse rewards. Experiments on the MIMIC-IV clinical dataset show that the proposed method learns more robust treatment policies and can decrease patient mortality by about 10.23% relative to baseline methods, demonstrating the power of reinforcement learning in personalised medical treatment planning.

In [12], the authors introduced an optimal control method that uses reinforcement learning to enhance the operation of power generation units, thereby improving operational efficiency. The integration of renewable energy, variable load demand, and the unpredictable disruption of the system are becoming a greater challenge to modern power systems. Conventional control measures are usually based on static optimisation models and a priori scheduling approaches that are not flexible enough to handle dynamic, uncertain environments. To address these limitations, the authors propose a deep reinforcement learning (DRL) framework that embeds the unit operation problem into the state-action-reward decision process. The system continuously monitors operational states and explores the environment to learn the best control actions. Another approach described in the study to optimise the system's performance is the application of Markov decision processes and more sophisticated optimisation methods, including the combination of reinforcement learning with algorithms such as the Hooke-Jeeves and deep deterministic policy gradient methods. The experimental analysis demonstrates that the proposed RL-based approach can enhance multi-objective optimisation, operational stability, and flexibility of power generation facilities compared with conventional control methods.

In [13], the authors developed an event-based control of drug dosage for tumour-immune systems using safe integral reinforcement learning (IRL). The study aims to enhance the efficacy of cancer immunotherapy by controlling drug intake without compromising system safety and stability. The classical forms of continuous control require frequent updates to the system and accurate models, which can raise the cost of computation and medication consumption. To address these constraints, the authors present an event-driven feature that changes the drug dosage only when certain system conditions are not met, thereby minimising unnecessary system control steps. The proposed method combines safe integral reinforcement learning to discover optimal control policies without requiring knowledge of the system dynamics. Moreover, safety limits are introduced to ensure that drug dosage is controlled and the immune system is not destabilised. Simulation experiments on nonlinear tumour-immune interaction models demonstrate that the proposed method successfully inhibits tumour growth, reduces drug use, and achieves better computational performance than conventional continuous control methods.

In [14], the authors examined the use of deep reinforcement learning (DRL) to optimise chemotherapy in a dynamical tumour growth model. The research question is the difficulty of identifying the optimal chemotherapy schedules that effectively suppress tumour cells with minimal adverse effects. The conventional methods of chemotherapy planning adopted are based on fixed treatment schedules or mathematical optimisation models that, in most cases, are inefficient at responding to dynamic tumour behaviours and patient-specific circumstances. To address these weaknesses, the authors develop a DRL-based control framework in which tumour growth and chemotherapy response are modelled as a sequence of decisions. The agent, based on reinforcement learning, interacts with the tumour model and learns optimal drug administration techniques by leveraging the rewarding nature of learning. The model will minimise drug overdose and maximise the reduction of tumour size by continuously changing the treatment response depending on the state of the system. The simulation results indicate that the proposed DRL method can significantly enhance treatment efficiency, stabilise tumour dynamics, and achieve better tumour suppression than traditional chemotherapy scheduling techniques.

In [15], the authors introduced an optimal dosage control scheme for the tumour-immune system using reinforcement learning (RL). The study aims to enhance cancer

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

immunotherapy by identifying an adaptive, optimal drug administration policy that controls the interaction between tumour cells and immune cells. Conventional control methods typically rely on accurate mathematical models and predetermined control laws, which can be ineffective in the presence of biological uncertainty and patient variability. To address this problem, the authors cast tumour-immune regulation as an optimal control problem based on reinforcement learning. The RL agent finds the optimal drug dosage strategy via repeated interactions with the tumour-immune dynamic model, minimising a cost that includes tumour growth and drug toxicity. The presented framework enables the controller to be trained on the best policies without necessarily knowing the system dynamics. The simulations show that the RL-based approach is an effective method for inhibiting tumour development, enhancing immune response control, and reducing drug overdose, making it appropriate for intelligent, personalised cancer treatment planning.

III. PROPOSED MODEL

The proposed system presents a DRL-ARDAC framework for controlling interactions within the tumor-immune system in cancer immunotherapy. The main aim of the proposed strategy is to identify an effective and safe dosage policy for drugs that can inhibit tumor growth without affecting immune cells or causing serious side effects from excessive drug use. Cancer immunotherapy involves complex biological mechanisms in which interactions between tumor and immune cells are nonlinear and unpredictable. These biological systems respond in different ways depending on factors such as patient-specific immune responses, tumor mutation rates, environmental disturbances, and treatment conditions. Classical model-based controllers rely on accurate mathematical models of biological interactions, but these models are not complete or accurate in practice in clinical settings. Therefore, the suggested system assumes a data-driven intelligent control policy based on deep reinforcement learning, which can learn optimal treatment policies by interacting with the system environment without precise knowledge of the system dynamics. The suggested DRL-ARDAC model incorporates three significant elements:

i. Modelling of tumor-immune dynamics.

ii. Control policy of deep reinforcement learning.

iii. Adaptive robust control under the bounded drug constraints.

The system constantly monitors the tumor-immune system's current state and determines the most appropriate drug dosage to achieve the greatest therapeutic impact while posing no risk to treatment. The process is performed with a critical neural network that approximates the optimal value function and updates the control policy through reinforcement learning. Moreover, a non-quadratic cost function with a discount is proposed to address the restriction on drug dosage and explicitly minimize toxicity.

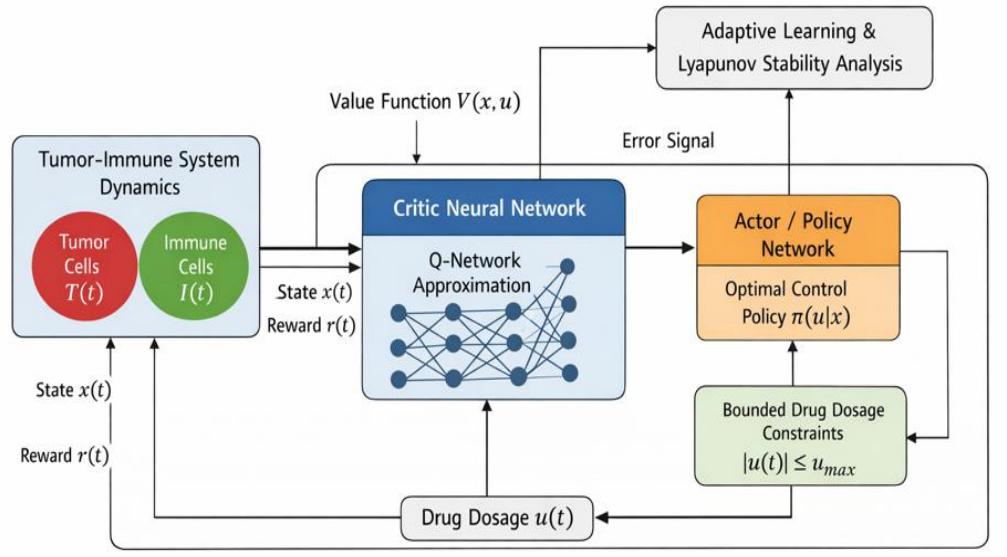

Fig.1. Framework for regulating tumor–immune system interactions

A. Tumour-Immune System Dynamic Model.

The tumor immune system dynamic model is a biological model of the interaction between tumour cells and immune cells, one of the vital regulatory models of cancer immunotherapy. The tumour-immune interaction in the proposed framework is modelled as a nonlinear dynamic system in which tumour cells proliferate and interact with immune effectors that destroy cancer cells. Let T(t) is the population of tumour cells, and $I(t)$ is the population of immune cells at time t. Natural proliferation and suppression of tumour cells by immune cells determine tumour development, and the presence of tumors and drug use stimulates immune cell activity. The drug dose u(t) is a control input that enhances the immune response and supports tumour suppression. The tumour dynamics are mathematically formulated as.

$$\frac{dT(t)}{dt} = aT(t) - bT(t)I(t) - u(t)T(t) \qquad (1)$$

with a being the rate of tumour growth, b being the rate of killing of tumour cells by the immune system, and $u(t)$, the dosage of drug administered. The following equation

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

represents the basic interplay among tumour growth, the immune response, and the therapeutic intervention. The fact is, however, that the precise parameter values are not readily available due to biological uncertainty and differences among individual patients. Hence, the suggested reinforcement learning control framework learns optimal drug dosage policies adaptively without requiring precise system modelling.

B. *Augmented State Representation*

To enhance the performance of the tumour-immune regulation system in terms of control, the proposed framework presents an augmented state representation that uses the biological system's states and the tracking error with respect to the therapeutic target. Physical states considered in conventional control systems are limited to tumour and immune cells. Nonetheless, to achieve effective cancer immunotherapy, one should also monitor the system's proximity to the intended treatment goal, which can be achieved by reducing tumour burden to an acceptable level and maintaining immune function. Thus, to construct an augmented state vector, the tumour population, the immune cell population and the tracking error are combined. Where $T(t)$ = tumour cell population, $I(t)$ = immune cell population, and e(t) = the error between the desired tumour cell population T_d and the actual tumour cell population $T(t)$. The augmented state vector has the form,

$$x(t) = [T(t), I(t), e(t)]^T \quad (2)$$

where the tracking error is defined as

$$e(t) = T_d - T(t) \quad (3)$$

This augmented-state formulation enables the reinforcement learning controller to learn both the system dynamics and the control goals simultaneously. The tracking error is incorporated into the state space, and the controller can learn drug dosage policies that ensure the desired tracking of the tumour population and achieve the intended therapeutic target without disrupting the stability of the immune system. The solution will help to increase treatment accuracy and strengthen the effectiveness of the offered framework of cancer immunotherapy control.

C. *Control Strategy-based on reinforcement learning.*

The proposed system will employ a RL control approach to identify the most effective drug dosage to control the dynamics of the tumour-immune system. Reinforcement learning enables an intelligent agent to learn optimal decision-making policies by interacting with an environment and receiving feedback in the form of rewards or penalties. In cancer immunotherapy, the RL agent monitors the tumour-immune system's current biological state and identifies a drug dose that inhibits tumour growth while preserving immune system stability. The control problem in tumour treatment is formulated as an MDP, comprising four major elements: state, action, reward, and policy. The state is the augmented system vector of the tumour cells, immune cells, and tracking error. This action is in line with the amount of drug that is given to the patient at a given time step. The reward mechanism assesses treatment efficacy by punishing large tumour populations, restraining excessive drug use within the cohort, and promoting a normal immune response. The RL agent gets an optimal control policy, which means system states to drug dosage actions. The goal of the learning process is to maximize the cumulative reward over time.

$$J = \sum_{t=0}^{\infty} \gamma^t r(t) \quad (4)$$

Where J is the cumulative reward, $r(t)$ is the immediate reward at time t, and γ is the discount factor ($0< \gamma <1$) that measures the value of future rewards. By continuously updating the control policy based on the tumour-immune system model, the reinforcement learning controller learns an optimal treatment strategy that minimizes tumour growth without endangering safe drug dosages.

D. *Critic Neural network architecture.*

Within the suggested DRL-ARDAC, the critic neural network assesses the quality of the control policy and estimates the long-term performance of the drug administration strategy. The critic network in reinforcement learning approximates the value function, which is the expected cumulative cost or reward for a given system state. The assessment supports learning and enhances the control policy governing tumour-immune dynamics. The augmented state vector $x(t)$, containing the tumour cell population T(t), the immune cell population I(t), and the tracking error $e(t)$, serves as the critic neural network's input. These inputs are fed to several hidden layers, which isolate nonlinear characteristics of the tumour-immune system's behaviour. The network has an output layer that approximates the value function, which determines the long-term effectiveness of the current drug dosage. The critic network estimates that the value function is of the form.

$$V(x) = W^T \phi(x) \quad (5)$$

Where $V(x)$ is the estimated value function, W is the weight vectors of the neural network, and $\phi(x)$ are nonlinear basis functions based on the state variables. Gradient descent is used to update the network weights to reduce the temporal difference error between the estimated and actual values. The critic network, through this learning process, provides accurate feedback on the treatment's effectiveness, enabling the reinforcement learning controller to determine better drug dosage and achieve optimal tumour suppression with minimal drug toxicity.

E. *Cost Function Design*

The cost function is critical to the proposed DRL-ARDAC structure because it helps steer the learning process toward the desired effective tumour suppression and the safety of drug administration. The cost function aims to punish unwanted system behaviors that include high tumour growth, huge tracking errors, and high levels of drug toxicity or undesirable side effects. With the inclusion of these factors, the reinforcement learning controller will be able to learn an optimal treatment strategy which would stabilize the therapeutic effectiveness and patient safety.

A discounted, non-quadratic cost function is constructed in the proposed system to directly impose limitations on drug dosage whilst reducing the number of tumour cells and stabilizing the immune system. The cost function may be put as

$$J = \int_0^{\infty} e^{-\gamma t} [Q_1 T^2(t) + Q_2 e^2(t) + R|u(t)|]dt \quad (6)$$

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

In which J is an accumulation of the treatment cost, $T(t)$ is the population of the tumour cell, $e(t)$ is the error rate in tracking the desired and actual tumour cell populations, and $u(t)$ is the dosage of the administered drug. Q_1 and Q_2 are the weighting parameters that determine the relative importance of tumour suppression and tracking accuracy, respectively, and R discourages drug overuse. The exponential discount factor $e^{-\gamma t}$is used to ensure that treatment outcomes are priority-driven while accounting for both immediate and long-term therapeutic performance. The reinforced learning agent can develop a drug administration policy using this cost formulation to minimize tumour growth, reduce drug toxicity, and stabilize immune system behavior.

F. Adaptive Robust Learning Mechanism.

The proposed DRL-ARDAC model consists of an adaptive, robust learning process to manage uncertainties and fluctuations in the interrelations between the tumour and the immune system. Biological systems are also complex and prone to various disturbances, including patient-specific immune responses, erratic tumour growth patterns, and systems that cannot be fully modelled. Conventional control methods frequently fail to hold the ground in such ambiguous situations. Thus, the proposed approach incorporates adaptive reinforcement learning to continually revise the drug dosage policy based on perceived system behavior. Under this mechanism, the reinforcement learning agent engages the tumour-immune system and monitors the current augmented state vector, which consists of the tumour cell population, the immune cell population, and the tracking error. It is on this information that the controller decides on the appropriate drug dosage. Once the control input is applied, the system's response is measured, and the learning algorithm adjusts the neural network parameters to enhance the control policy. This dynamic adaptation allows the system to respond to changes in biological conditions during treatment. The adaptive learning process can be expressed as

$$W_{k+1} = W_k + \alpha\delta_k\phi(x_k) \qquad (7)$$

with W_k being the weight vector of the neural network at iteration number k, a being the learning rate, δ_k being the temporal difference error, and $\phi(x_k)$ being the nonlinear feature of the system state. The update rule can be used to gradually improve the approximation of the value function and make the control policy more accurate. The adaptive robust learning mechanism also ensures that the control strategy performs well even in the presence of uncertainties, disturbances, and parameter changes through continuous system feedback. Consequently, the suggested framework will offer better robustness, stability, and flexibility of smart cancer immunotherapy treatment planning.

G. Lyapunov-Based Stability Analysis

A Lyapunov-based stability analysis is used to ensure the reliability and safety of the proposed DRL-ARDAC framework. The biomedical control systems need stability analysis, as uncontrolled behaviour or unstable behaviour can result in an overdose of the drug, immunosuppression, or uncontrolled tumour growth. Thus, the proposed control strategy aims to ensure that the tumour-immune system states remain within acceptable limits and to achieve the established therapeutic goal. The augmented system state vector is defined as $x(t) = [T(t), I(t), e(t)]^T$, where T(t) denotes the tumour cell population, $I(t)$ denotes the immune cell population, and $e(t)$ denotes the tracking error. The Lyapunov candidate function is designed to measure the stability of the closed-loop control system:

$$V(x) = \frac{1}{2}x^T Px \qquad (8)$$

where P is a positive-definite matrix in this case, and x is the augmented state. The Lyapunov function is the system's energy, which is always nonnegative except at the equilibrium point. To make the system stable, it is required that the time derivative of the Lyapunov function be negative definite.

$$\dot{V}(x) \leq -\lambda \parallel x \parallel^2, \quad \lambda > 0 \qquad (9)$$

Then the system's states are bounded, and the tracking error converges to zero. This state ensures that tumour-immune regulation is asymptotically stable. As a result, the suggested DRL-ARDAC controller will stabilize drug delivery while ensuring proper tumour removal and a non-hazardous immune response, even under conditions of uncertainty and interference.

IV. RESULTS AND DISCUSSIONS

To evaluate the feasibility of the proposed DRL-ARDAC framework, simulation experiments were conducted using nonlinear tumor-immune dynamic models. The simulations were performed on a system with an Intel Core i5 processor (3.4 GHz) and 8 GB of RAM. The results of the proposed approach were compared with three control methodologies that are typically used to control tumor-immune systems (i) Optimal Control Method (OCM) [16] (ii) Adaptive Control Method (ACM) [17] (iii) Normal Reinforcement Learning (RL) [18]. The assessment was based on key performance indicators such as tumour suppression percent, drug dosing efficiency, stability of immune response, and computation time. Experiments on simulations were performed under different conditions of tumour growth, including exponential and Gompertz growth models.

Table 1 compares the performance of various control strategies for tumour suppression. The findings reveal that the optimal control method has a tumour suppression rate of 82%, which cannot be extended to the dynamics of tumour progression because it relies on predetermined mathematical systems and is inflexible. The adaptive control method enhances the suppression rate to 86% because it can partially modify the control parameters according to the variations present in the system. The reinforcement learning method also extends the suppression performance to 90% because it acquires optimum treatment policies by interacting with the tumour–immune system environment. Nonetheless, the proposed DRL-ARDAC model has the best tumour suppression rate of 95%. This was achieved primarily because of the incorporation of deep reinforcement learning, augmented state representation, and critic neural network architecture, which allows the optimal drug dose to be predicted precisely. Moreover, the adaptive learning mechanism and constraint on drugs enable the system to control tumour growth without disrupting the stability of the immune system.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

TABLE 1: Tumor Suppression Performance

Method	Tumor Suppression Rate (%)
Optimal Control Method	82
Adaptive Control Method	86
Reinforcement Learning	90
Proposed DRL-ARDAC	**95**

Table 2 compares the average drug dosage consumption of various control approaches. The outcomes show that the optimal control method uses the highest drug dosage (0.78 units) partly because it uses fixed control strategies that are not very efficient in managing dynamic behavior between the tumour and immune system. The adaptive control method achieves the same with the drug dosage being reduced to 0.72 units by varying control parameters based on the variation in the system. In the same note, reinforcement learning also reduces the drug dose to 0.65, as it learns better drug administration methods as it becomes exposed to the system environment. Nevertheless, the proposed DRL-ARDAC framework has the lowest drug dosage consumption (0.58 units). This is enhanced by the incorporation of deep reinforcement learning, critic neural network, and constrained drug limits that allow the system to identify the most efficient dosage level. Consequently, the proposed approach reduces drug toxicity and ensures successful tumour repression and stability of the immune system.

TABLE 2: Drug Dosage Consumption Comparison

Method	Average Drug Dosage (Units)
Optimal Control Method	0.78
Adaptive Control Method	0.72
Reinforcement Learning	0.65
Proposed DRL-ARDAC	**0.58**

A comparison of the stability of the immune system with various control strategies is presented in Table 3. Immune system stability refers to the fact that the treatment plan can control the levels of immune cells within the normal physiological range and prevent tumour growth. The optimal control method attains an immune stability of 80%, which is a rather low score because the method is based on predetermined mathematical models that are incapable of adapting to biological uncertainties in all forms and situations. The adaptive control method enhances stability to 85%, where the control parameters are adjusted to meet changes in the system. Similarly, the reinforcement learning method is stable at a rate of 88% because it acquires optimal policies by being exposed to the tumor-immune system environment. Nevertheless, the proposed DRL-ARDAC framework attains the greatest immune stability of 93%. The augmented state representation, critic neural networks, and adaptive learning mechanisms are the reasons for this and allow one to carefully manage the dosage of the drug and maintain the balance of immune cells during treatment.

TABLE 3: Immune System Stability

Method	Immune Stability Score (%)
Optimal Control Method	80
Adaptive Control Method	85
Reinforcement Learning	88
Proposed DRL-ARDAC	**93**

Table 4 shows the levels of computational efficiency of various control methods. The average computation time, which is necessary to make optimal decisions in terms of drug dosage policies in the tumor immune system when regulating it, is used to measure computational efficiency. The optimal control method has the longest computation time of 2.4s because it uses the solution of complex mathematical equations of optimization repeatedly. The adaptive control method has a computation time of 2.1s because the control parameters are adjusted dynamically, although control parameter estimation is also necessary. The reinforcement learning method is also more efficient, with a calculation time of 1.8s, because it learns control policies as a result of interaction with the system environment. Nevertheless, the DRL-ARDAC framework suggested has the shortest computation time, 1.5s. This has been enhanced primarily by the critic neural network structure and adaptive learning process, which are effective at short-circuiting the optimal control policy and reducing the number of iterations needed to converge.

TABLE 4: Computational efficiency

Method	Computation Time (seconds)
Optimal Control Method	2.4
Adaptive Control Method	2.1
Reinforcement Learning	1.8
Proposed DRL-ARDAC	**1.5**

V. CONCLUSION

This study proposes the DRL-ARDAC framework for adaptive robust drug administration control in cancer immunotherapy. Unlike conventional model-based control strategies that rely on fixed mathematical models and predetermined parameters, the proposed framework employs deep reinforcement learning to learn optimal drug dosage policies through interaction with a nonlinear tumor-immune dynamic model. The augmented state representation, critic neural network, and non-quadratic discounted cost function collectively enable precise tumor suppression while enforcing drug dosage constraints and preserving immune system stability. Simulation results demonstrate that the proposed DRL-ARDAC framework outperforms existing control methods, achieving a tumor suppression rate of 95%, the lowest average drug dosage of 0.58 units, an immune stability score of 93%, and a computation time of 1.5 s. These results confirm that the proposed framework is an effective, robust, and computationally efficient solution for intelligent and personalized cancer immunotherapy. Future work will explore validation on clinical patient data, integration with

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

real-time biosensors, and extension to multi-drug treatment scenarios.

REFERENCES

[1] D. Hanahan and R. A. Weinberg, "Hallmarks of cancer: The next generation," Cell, vol. 144, no. 5, pp. 646–674, 2011. doi: 10.1016/j.cell.2011.02.013

[2] D. Kirschner and J. C. Panetta, "Modeling immunotherapy of the tumor–immune interaction," Journal of Mathematical Biology, vol. 37, no. 3, pp. 235–252, 1998. doi: 10.1007/s002850050130

[3] K. R. Fister and J. C. Panetta, "Optimal control applied to competing chemotherapeutic cell-kill strategies," SIAM Journal on Applied Mathematics, vol. 63, no. 6, pp. 1954–1971, 2003. doi: 10.1137/S0036139902405792

[4] J. Couzin-Frankel, "Cancer immunotherapy," Science, vol. 342, no. 6165, pp. 1432–1433, 2013. doi: 10.1126/science.342.6165.1432

[5] D. P. Bertsekas, "Dynamic programming and optimal control," Athena Scientific, 2012. doi: 10.1137/1.9781886529434

[6] V. Mnih et al., "Human-level control through deep reinforcement learning," Nature, vol. 518, pp. 529–533, 2015. doi: 10.1038/nature14236

[7] R. S. Sutton and A. G. Barto, "Reinforcement learning: An introduction," MIT Press, 2018. doi: 10.7551/mitpress/11447.001.0001

[8] F. Gottesman et al., "Guidelines for reinforcement learning in healthcare," Nature Medicine, vol. 25, pp. 16–18, 2019. doi: 10.1038/s41591-018-0310-5

[9] D. Liu and Q. Wei, "Adaptive dynamic programming for optimal control of nonlinear systems," IEEE Transactions on Neural Networks and Learning Systems, vol. 23, no. 6, pp. 913–924, 2012. doi: 10.1109/TNNLS.2012.2188316

[10] J. Hodgkin and A. Huxley, "A quantitative description of membrane current and its application to biological systems," Bulletin of Mathematical Biology, vol. 52, no. 1, pp. 25–71, 1990. doi: 10.1007/BF02464404.

[11] Q. Zhang, T. Li, D. Li, and W. Lu, "A goal-oriented reinforcement learning for optimal drug dosage control," Annals of Operations Research, vol. 338, no. 2–3, pp. 1403–1423, May 2024.

[12] G. Luo, T. Shi, C. Zhang, J. Jiang, and H. Li, "Optimal control strategy for unit operation based on reinforcement learning," Journal of World Architecture, vol. 9, no. 6, pp. 126–131, Dec. 2025.

[13] L. Chen, Y. Zhang, P. Yang, and X. Jin, "Event-triggered drug dosage control strategy of immune systems via safe integral reinforcement learning," European Journal of Control, vol. 82, p. 101201, Feb. 2025.

[14] W. Wang, L. Xu, D. Luo, J. Wu, and X. Wang, "Chemotherapy simulation of dynamical tumor model based on a deep reinforcement learning algorithm," Physica Scripta, Aug. 2025.

[15] L. Chen, Y.-W. Zhang, and S.-C. Zhang, "Optimal drug dosage control strategy of immune systems using reinforcement learning," IEEE Access, vol. 11, pp. 1269–1279, Jan. 2023.

[16] K. R. Fister and J. C. Panetta, "Optimal control applied to competing chemotherapeutic cell-kill strategies," SIAM Journal on Applied Mathematics, vol. 63, no. 6, pp. 1954–1971, 2003. doi: 10.1137/S0036139902405792

[17] D. Liu and Q. Wei, "Adaptive dynamic programming for optimal control of nonlinear systems," IEEE Transactions on Neural Networks and Learning Systems, vol. 23, no. 6, pp. 913–924, 2012.

[18] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)