Enhancing Cancer Prediction Accuracy Through Real-Time Monitoring and Artificial Intelligence Analysis for Patients

Zichang Wang *

Department of Artificial Intelligence and Advanced Computing, Xi'an Jiaotong-Liverpool University, Suzhou, China

* Corresponding Author Email: Zichang.Wang22@student.xjtlu.edu.cn

Abstract. The integration of Artificial Intelligence (AI) into clinical medicine has recently garnered significant attention, particularly in the context of digital pathology and precision medicine. Real-time monitoring and analysis of patients play a pivotal role in the overall treatment process. This includes monitoring basic body detection indices, analyzing the patient's overall condition, and predicting future possibilities for their medical status. In addition to fundamental machine learning (ML), both deep learning (DL) and digital twin technology have demonstrated considerable potential in enhancing the accuracy of analysis and prediction in clinical cancer patients' condition monitoring. This study employs a comparative analysis approach to assess and compare how DL and digital twin technology contribute to improving accuracy in monitoring and analyzing clinical cancer patients' conditions. ML is well-suited for training models with large datasets, while DL excels in tackling more intricate problems. To enhance the precision and accuracy of analysis and future prediction, digital twin technology is leveraged to explore various patient scenarios, with basic twins providing real-time monitoring of the patient's current conditions. Markov Decision Progress (MDP) is employed when predicting potential outcomes. However, it's essential to consider the impact of the number of times the model is trained when making comparisons. The evaluation criterion centers on whether the use of DL and digital twin technology improves accuracy and precision compared to using ML alone. Examining the results reveals that using only ML or DL yields accuracies of 97.01% and 94.73%, respectively. However, when ML is combined with DL and digital twin technology, accuracy significantly improves to 99.73%.

Keywords: Deep learning; digital twin; precision oncology; digital pathology.

1. Introduction

The rapid advancement of AI, particularly in ML and DL, alongside progress in digital twin technology and GPUs, coupled with the development of public databases, has paved the way for the integration of these technologies in the clinical treatment of cancer patients [1,2]. This integration has garnered global interest in artificial intelligence [3]. Unlike the previous two AI booms, this third wave is not primarily focused on social applications but has found utility in domains like airports. AI-equipped medical devices are already being used to treat clinical cancer patients, particularly in precision oncology. As digital twin technology emerges and public databases are established, the trend of using AI in clinical settings, including oncology, is poised for significant growth. In oncology, AI has demonstrated benefits, particularly in radiology, where deep learning and graphical neural networks aid in patient condition analysis [4].

Machine learning is currently the core AI technology. ML, in a nutshell, involves using machines to learn from historical data and formulate policies to analyze new data and solve specific problems. It can be categorized into supervised and unsupervised learning. Supervised learning recognizes patterns in previous data to address classification and regression problems, such as distinguishing between benign and malignant tumors, and making predictions, or comparisons—thus, it has found wide application in radiology [5]. Unsupervised learning, while less applicable in oncology, relies on prior data to guide machine learning in a specialized manner. The use of fundamental ML technology is integral to AI's clinical application. However, for more precise analysis and prediction of patient conditions, a more extensive dataset is required [6]. This is where DL and digital twin technology come into play. By creating digital twins of patients to gather health data, doctors can generate

additional twins for different patients, with each twin representing a potential future condition. Twins can analyze current data and forecast various scenarios using the MDP [7]. In oncology, gene analysis is crucial, and ML, DL, and digital twin technology significantly contribute to achieving digital pathology. This paper outlines the history and present use of AI in clinical cancer treatment and proposes a novel approach to enhance precision by creating multiple digital twins to explore diverse possibilities.

2. Related Work

This section aims to provide an explanation and detailed overview of the current methodologies. It also highlights some of the existing limitations and disadvantages. It's worth noting that the intricacies and comprehensive nature of oncology make it well-suited for AI-driven patient data monitoring and analysis, significantly benefiting cancer treatment. The predominant technologies employed in precision oncology and digital pathology today primarily revolve around machine learning, including deep learning. While digital twin technology has yet to find widespread use in oncology, it has begun to make inroads into clinical treatment [8]. To manage vast amounts of patient data, machine learning offers an avenue for analysis and prediction based on historical data. Essentially, when machines have learned from extensive prior data and established algorithmic systems to address classical problems, they can autonomously provide analyses and treatment solutions. This enables doctors to allocate more time to tackle the more intricate, non-automatable aspects of their profession [9]. However, when dealing with more intricate or specialized issues, basic machine learning falls short. For precise analysis and prediction of complex or specialized problems, DL is highly recommended. For instance, in tasks like image recognition, such as discerning the presence of tumors in images or determining if a tumor is benign or malignant, DL excels. This aspect is pivotal in the application of AI in precision oncology and digital pathology, particularly in conjunction with radiology [10]. By deeply analyzing images from various dimensions, machines can calculate a tumor's size and compare it with previous data to provide an accurate growth rate, thus predicting its future progression. Additionally, DL can extend beyond mere recognition. By analyzing tumor positions within a patient's body, it can compute various rates, including tumor transfer and progression rates.

However, despite the usefulness of DL in offering suggestions, it has a drawback: limited explanatory capability. In some cases where the treatment decision involves both the machine and the doctor, if the final results are unsatisfactory, it becomes challenging to pinpoint the source of errors. Yet, if DL is employed to enhance the precision and accuracy of monitoring, analysis, and prediction, extensive prior data is essential [11]. Consequently, when connected to extensive public databases, DL's role in clinical treatment becomes significantly more substantial. Without a substantial data foundation, DL cannot fully exploit its advantages in precision enhancement. Besides DL, neural network technologies are also critical in the realm of AI. A key component utilized in clinics is the graphical neural network (GNN). Some mature solutions have emerged for specific conditions in the clinical treatment of cancer patients using GNN. These solutions, in combination with DL and radiology, particularly for image-related issues as mentioned earlier, have yielded promising results.

As AI continues to advance, digital twin technology has found applications across various domains, notably in manufacturing. Although its use in healthcare, particularly in the context of cancer patient treatment, is not yet widespread, it is in its infancy. Consequently, applying digital twin technology to the clinical treatment of cancer patients presents a significant challenge, given the inherent complexity of the human system compared to physical objects. Despite the numerous advantages demonstrated by AI in treating cancer patients, realizing real-time digital monitoring and analysis remains elusive. If digital twin technology can be successfully and seamlessly integrated into this scenario, it holds the potential to greatly enhance the role of both machine learning and DL. The

aspiration for achieving real-time monitoring, analysis, and even precision prediction of a patient's condition would no longer remain a mere dream.

3. Methodology

Different methods have been proposed to categorize the application of AI in medicine. Traditionally, AI's role in clinical treatment has been evaluated by quantifying its contribution to the entire therapeutic process, with a particular emphasis on deep learning. However, to enhance AI's role and accuracy in clinical treatment and achieve the digital pathology goals discussed earlier, this paper introduces a novel concept: real-time monitoring and analysis of cancer patients using AI. This approach combines digital twin technology with a database and leverages deep learning to significantly enhance precision based on the currently used machine learning technology. This section provides a comprehensive discussion of the methodology's steps and elucidates the components with detailed diagrams and formulas.

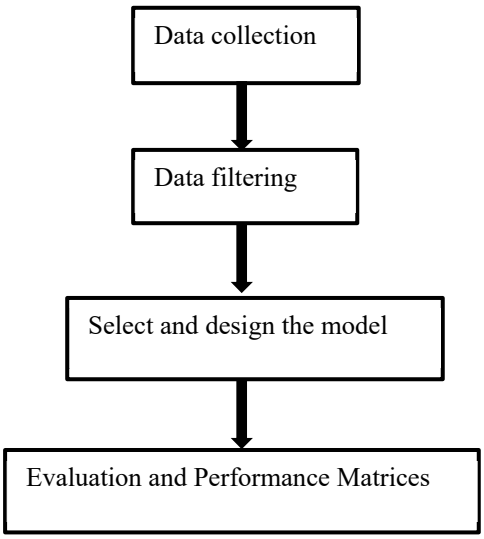


Fig. 1 The process of the method

Fig. 1 represents a simplified flowchart of the process. A more detailed closed-loop process diagram will be presented later in this section. The entire process of the new method for improving AI's accuracy and precision is outlined above. The initial step involves data collection and preparation. To ensure the data's freshness and abundance, all data used in this paper is sourced from the International Agency for Research on Cancer and the Journal of the National Cancer Center. Additionally, to test the future results of the final model, the data used for algorithm exploration and the establishment of the basic digital twin is not the most recent but rather from a short time before. One patient's data is randomly selected to represent the assumed patient undergoing therapy. Following the collection and preparation of a substantial volume of data, the subsequent step in the new method is data filtering. To effectively treat a cancer patient using AI, relevant data similar to their condition must be identified. This is followed by ML, GNN, and other processes. In this paper, data selection is executed using a precise filtering model, as it is not the primary focus of this research, and therefore, a detailed explanation is not provided.

After filtering the data, the next step in the method involves integrating digital twin technology with DL. To acquire more patient data, digital twin technology is employed to simulate future data scenarios, presenting different possibilities and enabling more precise predictions and treatments. However, building these digital twins and making accurate predictions requires several steps. Firstly, the filtered data should be employed for basic machine learning, including GNN for image analysis and other index learning algorithms. Consequently, the first model for fundamental ML and GNN is established. The model's core algorithm involves recognizing which specific situations are advantageous for the patient's current condition and which are not. Notably, the detailed information

should include image data, as GNN can analyze future tumor growth possibilities, such as size changes and the likelihood of metastasis. Ensuring the precision of the foundational training relies on considering only the most common features. Furthermore, to enhance the accuracy of the assumed patient's possibilities, the MDP is utilized. The comprehensive procedure is depicted in Fig. 2:

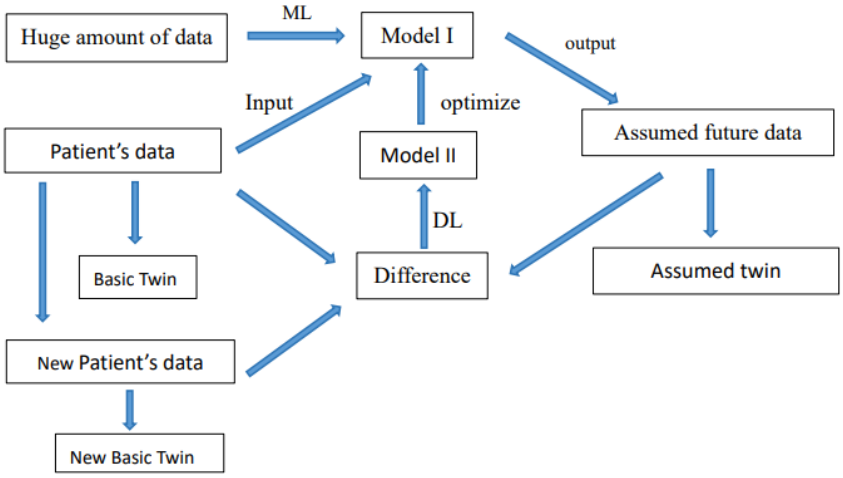


Fig. 2 The whole process of the method

Based on the patient's current physical characteristics, a foundational digital twin is established. Its purpose is to serve as a reference model for other twins representing potential future scenarios. Running the foundational twin may render some previously assumed possibilities invalid while introducing new ones. Consequently, the discrepancy between past assumptions and the current reality serves as the patient's personal data. This data is instrumental in creating and training the second model in this new method: the DL model. By employing this data difference with the DL model, prediction accuracy and precision are significantly enhanced. Additionally, digital twins can be employed to simulate treatment solutions. If successful, this represents a significant advancement in precision medicine and digital pathology. To further improve the analysis and prediction accuracy, two models are established. Model I provides initial predictions, and with updated patient data and the foundational twin, Model II learns from the differences between past predictions, current realities, and past and current data to optimize Model I's algorithms. This refinement results in Model I offering more accurate analyses and predictions in the near future. Formula 1 demonstrates the method for calculating probabilities. This formula is connected with the evaluation process; the closer this possibility is to 1, the higher the accuracy and precision of the model. With repeated utilization of reinforcement learning and Model II, the accuracy of Model I continually improves. In this method, digital twins replace previously assumed impossible scenarios with new possibilities, allowing these twins to explore all potential patient condition scenarios in the future. This results in a closed-loop process for achieving real-time analysis, monitoring, and prediction for patients.

$$Pr(Ai|B)=Pr(Ai)Pr(B|Ai)/\Sigma jPr(Aj)Pr(B|Aj) \quad (1)$$

GNN is primarily employed in image analysis. The underlying principles and functionalities of GNN in the context of clinical image-based problem-solving are outlined below. As depicted in the illustration, GNN can systematically monitor various states within clinical images and provide precise predictions through its inherent capabilities. This real-time monitoring of a patient's condition and treatment process becomes significantly streamlined with the incorporation of GNN. Furthermore, GNN can be seamlessly integrated with deep learning models to address specific clinical challenges with a high degree of effectiveness and accuracy. Fig. 3 depicts the graph encoding network.

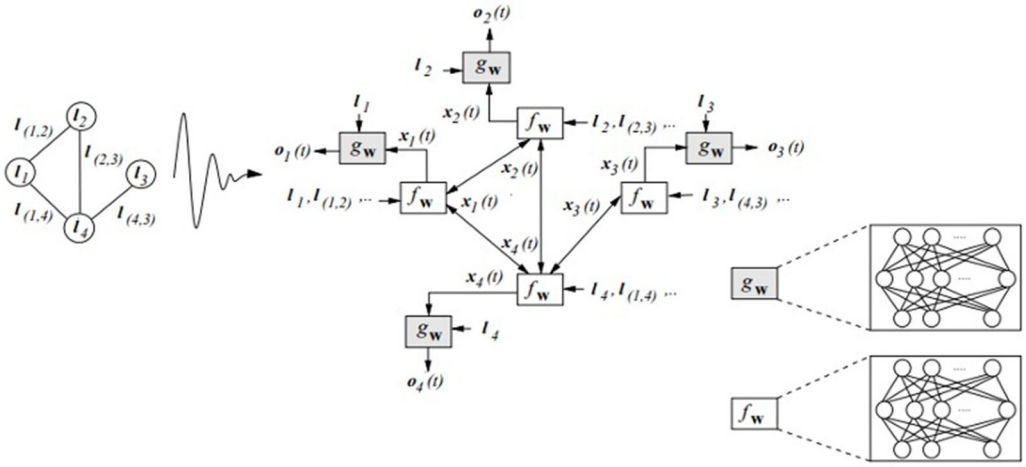


Fig. 3 The graph encoding network

4. Results and Analysis

Table 1 provides an overview of the computer hardware utilized in this experiment, while Table 2 summarizes the software employed.

Table 1. Hardware configuration

Component	Specification
Operating System	Windows 11
Memory	16GB
Processor	AMD R9-7945HX

Table 2. Software configuration

Library	Version
TensorFlow	2.12.0
NumPy	1.23.0
Keras	2.12.0
Pandas	1.5.3

To highlight the advantages of incorporating DL and digital twin technology over using ML alone, two sets of experiments have been designed. The key distinction between the two experiments lies in the utilization of digital twin technology. The dataset employed for the study is not entirely dedicated to model training; 70% of the data is allocated for model training, leaving the remaining 30% for assessing model accuracy. Within each group, three specific modeling approaches are implemented. Due to the complexity of the patient indices, the charts only indicate whether the predicted indices fall within the correct range. Detailed patient physical indices, such as mean radius, mean texture, mean perimeter, concave points worst, and others, are not presented here. The model accuracy results are detailed in Table 3.

Table 3. Experiment I (without digital twin tech)

Model/Index	Number of Right	Number of Wrong	Accuracy
C5.0 (ML)	552	17	97.01%
CNN (DL)	539	30	94.73%
C5.0+CNN(ML+DL)	555	14	97.58%

In this experiment, a total of 569 data points is used to assess model accuracy. The ML+DL model achieves the highest accuracy, surpassing the DL model, albeit marginally. This discrepancy may be

attributed to the relatively limited dataset size, which potentially restricts DL's ability to discover a more precise algorithm. To address this limitation, Table 4 presents Experiment II, which employs a larger dataset for model training. Here, the DL model demonstrates its accuracy advantage. However, it is worth noting that the ML+DL model consistently outperforms others in terms of prediction accuracy.

Table 4. Experiment II (with digital twin tech)

Model/Index	Number of Right	Number of Wrong	Accuracy
C5.0 (ML)	558	11	98.07%
CNN (DL)	565	4	99.30%
C5.0+CNN(ML+DL)	568	1	99.82%

The primary distinction between this experiment and the first lies in the inclusion of digital twin technology. With significantly larger training datasets, the models can explore more precise algorithms. Consequently, the DL model demonstrates an advantage in analyzing and predicting patient conditions. Nevertheless, it remains evident that the ML+DL model excels in both analysis and prediction.

Comparing the two sets of experiments, it becomes evident that digital twin technology significantly enhances model accuracy. Additionally, the DL model benefits from larger datasets for training. In scenarios with very limited data, the accuracy of the CNN model fails to exhibit any advantage, even when compared to the basic ML model.

5. Conclusion

This paper introduces a methodology for the treatment of cancer patients and the realization of real-time clinical monitoring and analysis using AI. The approach incorporates various technologies, including fundamental machine learning and deep learning, while also integrating digital twin technology. Two distinct models have been established, demonstrating advantages in accuracy and precision over a defined period of treatment when compared to the use of ML alone. However, it is important to note that the new method, while offering improvements in accuracy and precision, still exhibits certain limitations. In the future, there is a need for further optimization of these models in conjunction with digital twin technology. This represents a promising avenue for combining clinical treatment with AI technology, marking significant progress in the field of digital pathology and providing substantial support for precision oncology. Moreover, the methodology employed in this study holds the potential for personalized treatment. By tailoring the approach to individual patients, each individual can benefit from more precise and customized care, ultimately enhancing the quality of healthcare delivery. As AI continues to advance and as more data becomes available, the fusion of these technologies holds the promise of revolutionizing cancer treatment and improving patient outcomes.

References

[1] Shimizu H, Nakayama K I. Artificial intelligence in oncology [J]. Cancer science, 2020, 111(5): 1452-1460.

[2] Barricelli B R, Casiraghi E, Fogli D. A survey on digital twin: Definitions, characteristics, applications, and design implications [J]. IEEE access, 2019, 7: 167653-167671.

[3] Hamamoto R, Komatsu M, Takasawa K, et al. Epigenetics analysis and integrated analysis of multiomics data, including epigenetic data, using artificial intelligence in the era of precision medicine [J]. Biomolecules, 2019, 10(1): 62.

[4] Rudie J D, Rauschecker A M, Bryan R N, et al. Emerging applications of artificial intelligence in neuro-oncology [J]. Radiology, 2019, 290(3): 607-618.

[5] Hamamoto R, Suvarna K, Yamada M, et al. Application of artificial intelligence technology in oncology: Towards the establishment of precision medicine [J]. Cancers, 2020, 12(12): 3532.

[6] Chua I S, Gaziel-Yablowitz M, Korach Z T, et al. Artificial intelligence in oncology: Path to implementation [J]. Cancer Medicine, 2021, 10(12): 4138-4149.

[7] Kaul R, Ossai C, Forkan A R M, et al. The role of AI for developing digital twins in healthcare: The case of cancer care [J]. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2023, 13(1): e1480.

[8] Holzinger A, Langs G, Denk H, et al. Causability and explainability of artificial intelligence in medicine [J]. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2019, 9(4): e1312.

[9] Iqbal M J, Javed Z, Sadia H, et al. Clinical applications of artificial intelligence and machine learning in cancer diagnosis: looking into the future [J]. Cancer cell international, 2021, 21(1): 1-11.

[10] Abdel Razek A A K, Alksas A, Shehata M, et al. Clinical applications of artificial intelligence and radiomics in neuro-oncology imaging [J]. Insights into imaging, 2021, 12: 1-17.

[11] Eckardt J N, Wendt K, Bornhaeuser M, et al. Reinforcement learning for precision oncology [J]. Cancers, 2021, 13(18): 4624.