

Artificial intelligence empowered optimization of tumor single-cell sequencing data features and clinical implications

Yuxia Yang^a, Tu Ya^a, Chao Luomeng^{*b,c,d}

^aCollege of Computer Science and Technology, Inner Mongolia Minzu University, Tongliao 028000, Inner Mongolia, China; ^bCollege of Animal Science and Technology, Inner Mongolia Minzu University, Tongliao 028000, Inner Mongolia, China; ^cInner Mongolia Key Laboratory of Toxicant Monitoring and Toxicology, Tongliao 028000, Inner Mongolia, China; ^dRambo Animal Disease Inspection and Quarantine Animal Disease Inspection and Quarantine Limited Company, Tongliao 028000, Inner Mongolia, China

ABSTRACT

With the rapid development of personalized medicine and precision treatment, tumor single-cell sequencing technology has highlighted its important role in the study of tumor heterogeneity. This article systematically reviews the technological progress and current applications of single-cell sequencing technology in the field of tumors, and explores in depth how artificial intelligence (AI) can promote feature optimization of single-cell data and its potential value in tumor diagnosis and treatment. This study introduces the current mainstream data optimization methods and analyzes how artificial intelligence technology can improve diagnosis and treatment strategies through case studies, pointing out the advantages of artificial intelligence in improving the accuracy of single-cell sequencing, especially in enhancing tumor biomarker identification, predicting recurrence potential, and assisting personalized treatment decision-making. This article concludes that artificial intelligence technology has broad prospects in optimizing precise tumor diagnosis and treatment, effectively promoting the transformation of tumor management, and providing new ideas for the innovation of future tumor treatment models.

Keywords: Single-cell sequencing of tumors, artificial intelligence, data feature optimization, precision diagnosis and treatment, heterogeneity analysis, treatment strategy

1. INTRODUCTION

The application of artificial intelligence (AI) in the medical field is becoming increasingly widespread, particularly in tumor diagnosis and treatment. Optimizing the features of tumor single-cell sequencing data holds significant promise for enhancing tumor diagnosis and treatment outcomes. This article explores the optimization of tumor single-cell sequencing data features and its impact on diagnosis and treatment, leveraging AI technology. Tumor single-cell sequencing data contains vast amounts of information, and AI technology can uncover potential patterns and features within this massive data-set, providing more accurate foundations for tumor diagnosis¹.

Optimizing the features of tumor single-cell sequencing data can help physicians diagnose tumor types and assess the severity of conditions, while also supporting personalized treatment strategies. AI algorithms can predict the development and metastasis of tumors based on the genomics characteristics of patients, offering valuable insights for doctors to formulate individualized treatment plans. Moreover, AI can perform rapid and efficient analyses of tumor single-cell sequencing data, significantly reducing the time required for diagnosis and treatment. Additionally, AI technology can assist medical researchers in discovering new tumor features, providing innovative approaches for early diagnosis and prevention of tumors².

AI-assisted feature optimization of tumor single-cell sequencing data promises revolutionary changes in tumor diagnosis and treatment. Through AI technology, tumor single-cell sequencing data can be analyzed more accurately and swiftly, offering patients more precise diagnoses and personalized treatment plans, thereby improving tumor cure rates and survival rates.

*chuoluomeng@imun.edu.cn

2. OVERVIEW OF SINGLE-CELL SEQUENCING IN TUMOR RESEARCH

2.1 Development of single-cell sequencing technology

The development of single-cell sequencing technology is rapid and transformative, becoming an essential tool in tumor research. From the initial single-cell RNA sequencing and DNA sequencing to current advancements in proteomics sequencing, significant progress has been made. As single-cell technologies continue to innovate, our understanding of tumor cell heterogeneity and potential mutations will deepen, driving progress in individualized tumor treatment. AI plays a crucial role in this technological advancement. Through computer models and algorithms, AI can analyze and mine large amounts of single-cell sequencing data, discovering potential tumor characteristics and clinical significance. These technological methods not only enhance the efficiency of data interpretation but also identify features that may be invisible to the human eye³.

Single-cell sequencing technology is powerful in tumor diagnosis and treatment. By sequencing individual tumor cells, we gain a clearer understanding of tumor heterogeneity, providing a more accurate basis for precision therapy. Single-cell sequencing can also help doctors identify drug-resistant cells within a patient, enabling adjustments to treatment plans and improving treatment effectiveness.

In the future, single-cell sequencing technology will play an increasingly important role in tumor research and clinical applications. Intelligent AI analysis will further aid researchers in discovering and understanding tumor cell characteristics, providing more precise recommendations for personalized treatment. Single-cell sequencing technology is poised to become an invaluable tool in tumor research, leading to better diagnostic and therapeutic outcomes for tumor patients.

2.2 Current applications in tumor research

The application of single-cell sequencing technology in tumor research has become a hot topic. Single-cell sequencing allows for a comprehensive understanding of tumor cell heterogeneity, the discovery of potential therapeutic targets, and crucial support for personalized treatment and targeted drug development. AI, as an emerging technology, is gradually permeating various aspects of single-cell sequencing data analysis in tumor research, offering new methods for optimizing the characteristics of tumor single-cell sequencing data.

The application of AI in optimizing tumor single-cell sequencing data features can help identify and uncover hidden biological characteristics, enhancing the efficiency and accuracy of data interpretation. Using AI algorithms to analyze large-scale tumor single-cell sequencing data can identify new tumor cell subgroups and discover potential targets for targeted therapy, providing comprehensive information for tumor diagnosis and treatment. In breast cancer research, clustering algorithms like t-SNE and UMAP, and deep learning models such as autoencoders, efficiently process and analyze single-cell RNA sequencing data, allowing researchers to identify new tumor cell subgroups in highly heterogeneous tumor tissues⁴.

AI can also perform pattern recognition and feature extraction on tumor single-cell sequencing data, helping researchers better understand the interactions and signaling pathways between tumor cells. For example, melanoma, a highly malignant skin cancer, involves complex cellular interactions and signaling pathways in its progression and metastasis. Traditional research methods struggle to comprehensively reveal these complex biological processes. Advanced AI algorithms, like Graph Neural Networks (GNNs) and Self-Attention Mechanisms, can analyze single-cell RNA sequencing data of melanoma; GNNs handle and analyze relationships between cells, while self-attention mechanisms capture important patterns and features in the data. AI algorithms can discover patterns and trends from vast amounts of data, providing new perspectives and interpretations for tumor research⁵.

In tumor research, AI can also accelerate new drug development and clinical translation. Using AI algorithms to perform network modeling and target prediction on tumor single-cell sequencing data enables drug developers to quickly screen potential targeted drugs, assess their efficacy and side effects in advance, and provide a more scientific basis for clinical treatment. Lung cancer, a leading cause of cancer death worldwide, exemplifies the high heterogeneity challenge where many patients do not achieve ideal treatment outcomes. New drug development requires significant time and resources, but AI algorithms can model networks and predict targets based on single-cell RNA sequencing data from lung cancer patients, accelerating new drug development and clinical translation. The application of AI expands the depth and breadth of tumor research, presenting new opportunities and challenges for personalized treatment and targeted drug development.

3. APPLICATION OF ARTIFICIAL INTELLIGENCE IN DATA FEATURE OPTIMIZATION

3.1 Theoretical foundation of optimization methods

In the process of optimizing features of tumor single-cell sequencing data, AI technology plays a crucial role in data analysis and processing, achieving feature optimization and providing more precise guidance for tumor diagnosis and treatment. The application of AI in data feature optimization must be built on a solid theoretical foundation, which includes knowledge in data processing, feature extraction, and model training.

In terms of data processing, AI technology can preprocess raw data through methods such as data cleaning, denoising, and normalization, laying the groundwork for subsequent feature extraction and model training. For feature extraction, AI employs various algorithms and models to identify key features within the data, thus optimizing data characteristics. As for model training, AI technology leverages deep learning and machine learning techniques to train and refine the extracted data features, delivering more accurate results for tumor diagnosis and treatment.

The theoretical foundation of optimization methods also encompasses in-depth research on the application of AI in optimizing tumor single-cell sequencing data features. This involves analyzing the principles and applicable scopes of various optimization methods, evaluating the effectiveness of different methods in practical applications, and exploring future trends and directions. A robust and comprehensive theoretical foundation enables AI technology to demonstrate significant value in optimizing tumor single-cell sequencing data features. Fully understanding and grasping this theoretical foundation allows for better utilization of AI technology to enhance tumor single-cell sequencing data features, providing more precise support for tumor diagnosis and treatment.

3.2 Analysis of practical application cases

As an emerging technology, artificial intelligence is widely applied in optimizing tumor single-cell sequencing data features. In practical applications, AI can quickly and accurately identify key features from large datasets through learning and analysis, offering more precise support for tumor diagnosis and treatment. By examining practical application cases, we can gain insights into the effectiveness and potential of AI in this domain.

In recent years, AI has demonstrated numerous successful cases in the optimization of tumor single-cell sequencing data features. As shown in Table 1, For example, in tumor diagnosis, AI can swiftly and accurately screen potential pathogenic genes from patients' single-cell sequencing data, providing clinicians with more reliable diagnostic evidence. An AI model successfully identified several known and newly discovered pathogenic genes from the single-cell sequencing data of breast cancer patients. These genes include classic breast cancer-associated genes (such as BRCA1 and BRCA2) and emerging regulatory genes. The model's accuracy significantly surpasses traditional methods, offering rapid and precise diagnostic evidence for clinicians⁶⁻⁸.

Table 1. Applications of artificial intelligence in tumor single-cell sequencing.

Application scenario	Specific application direction
Tumor Prediction and Diagnosis	AI models can quickly and accurately screen for potential pathogenic genes, providing a diagnostic basis.
Personalized treatment plans	Targeted therapy and immunotherapy targets, assisting clinical doctors in diagnosis and treatment planning.
Intelligent image recognition system	X-rays, CT scans, MRI, pathology slides, etc.
Drug development	Target screening, drug screening, small molecule drug crystal structure prediction.
Health management	AI wearable medical devices, tumor monitoring, tumor lesion tracking, health guidance.
Tumor medical informatization	Establishment of intelligent tumor disease databases, data retrieval, model prediction, early cancer screening, risk warning.

In tumor treatment, AI can also provide personalized treatment plans based on the optimized features of single-cell sequencing data, enhancing the precision and effectiveness of treatment. For instance, an AI model successfully

identified suitable targets for targeted therapy and immunotherapy from the single-cell sequencing data of lung cancer patients. For certain patients, the targeted drugs and immunotherapy drugs recommended by the model significantly improved treatment outcomes, extending both progression-free survival (PFS) and overall survival (OS)⁹.

AI's application in image and pathology analysis has reached a relatively mature level of speed and accuracy. Cicero et al. employed CNN deep learning on 35,038 chest X-rays to label pleural effusion, cardiomegaly, lung consolidation, pulmonary edema, and pneumothorax, achieving high specificity and sensitivity, up to 91%. Esteva et al. trained a CNN-based AI system to identify skin cancer from nearly 130,000 images of moles, rashes, and other skin lesions, achieving an accuracy rate of over 91%, comparable to that of 21 dermatologists. Hazlett et al. utilized a 3-layer deep learning network to reduce the dimensions of 315 parameters (including gender, volumes, areas, and thicknesses of different brain regions) to predict the likelihood of autism in high-risk infants, achieving an accuracy of 94%. Additionally, in the grading and classification of brain tumors, prostate cancer, lung nodules, and breast cancer, accuracies of around 90% have been achieved.

Through learning and analyzing large amounts of data, AI can quickly and accurately identify key features, providing more precise diagnostic and treatment support for clinicians. In this study, three classification networks—Graph Convolutional Neural Network (G-CNN), Visual Geometry Group 19 (VGG19), and Residual Network 50 (ResNet50)—were optimized for tumor region classification. As shown in Table 2, the accuracy of these three models improved after optimization, particularly in the G-CNN and ResNet50 models. The performance of different models varied across different grades.

Table 2. Accuracy rates of different grade levels in tumor regions before and after network optimization.

Classification network		3rd Grade	4A Grade	4B Grade	4C Grade	5th Grade
Graph convolutional neural network (G-CNN)	Before optimization	0.956 (0.0077)	0.898 (0.0113)	0.680 (0.0679)	0.907 (0.0435)	0.838 (0.1233)
	After optimization	0.999 (0.0041)	0.941 (0.0111)	0.735 (0.0663)	0.923 (0.0377)	0.877 (0.1235)
Visual geometry group 19 (VGG19)	Before optimization	0.948 (0.0146)	0.865 (0.0133)	0.668 (0.0727)	0.866 (0.0464)	0.819 (0.1282)
	After optimization	0.991 (0.0048)	0.921 (0.0195)	0.674 (0.0693)	0.909 (0.0494)	0.842 (0.1320)
Residual network 50 (ResNet50)	Before optimization	0.955 (0.0137)	0.879 (0.0125)	0.670 (0.0665)	0.900 (0.0426)	0.836 (0.1339)
	After optimization	0.992 (0.0053)	0.928 (0.0156)	0.689 (0.0666)	0.921 (0.0389)	0.859 (0.1424)

In grades 3 and 4A, all models exhibited very high accuracy both before and after optimization, with the G-CNN's grade 3 accuracy reaching 0.999 after optimization. The improvement in grade 4B was less significant than in other grades, possibly due to the complexity of samples or features in grade 4B. Nevertheless, the G-CNN still achieved an accuracy of 0.735 after optimization. In grades 4C and 5, the accuracy of all models improved post-optimization, with the G-CNN showing the best performance in grade 4C. The G-CNN demonstrated the most significant overall improvement across all grades, making it the best choice for classifying tumor regions. The improvement of VGG19 was relatively small, especially performing worse than the other two models in grade 5. ResNet50's performance was stable overall after optimization, but slightly lower than G-CNN in grades 3 and 4A. In summary, the G-CNN showed the highest classification accuracy after optimization in most grades, making it the most superior model, recommended for priority consideration in practical applications.

Additionally, AI algorithms can analyze multimodal data of breast cancer patients (such as imaging data, genomic data, and pathology data) to extract key features, assisting clinicians in diagnosis and treatment planning. AI models, through learning and analyzing multimodal breast cancer data, successfully extracted key diagnostic and therapeutic features, significantly improving diagnostic accuracy and treatment effectiveness. This provides better treatment opportunities for

cancer patients and more reliable decision-making support for clinicians.

The analysis of practical application cases demonstrates that AI can swiftly and accurately identify key features in data by learning from and analyzing large datasets. This provides more precise support for the diagnosis and treatment of tumors, bringing new hope for tumor patients and offering greater convenience and support for clinicians. It is believed that with the continuous development and improvement of AI technology, the feature optimization of tumor single-cell sequencing data will make greater progress, bringing more positive impacts on the diagnosis and treatment of tumors.

4. THE ROLE OF ARTIFICIAL INTELLIGENCE IN DIAGNOSTIC AND THERAPEUTIC IMPACT

4.1 Improvement of diagnostic processes

In modern medical clinical diagnosis and treatment, early diagnosis and treatment of tumors are critical steps. With the continuous advancement of technology and the application of AI, the diagnostic process of tumors has been significantly improved and optimized. AI technology plays a crucial role in the optimization of tumor single-cell sequencing data features and its impact on diagnosis and treatment.

In the past, tumor diagnosis often relied heavily on the experience and expertise of doctors, occasionally resulting in missed diagnoses or misdiagnoses. However, with the introduction of AI, the diagnostic process has undergone revolutionary improvements. AI can rapidly and accurately identify tumor features and variations by analyzing and learning from large amounts of tumor single-cell sequencing data, thereby improving the accuracy of tumor diagnosis and early detection rates. The application of AI technology not only enhances the diagnostic process but also improves the work efficiency of doctors. Traditionally, doctors needed to spend significant time and effort analyzing tumor single-cell sequencing data to make diagnoses and treatment plans. With AI technology, doctors can quickly obtain accurate diagnostic results and formulate personalized treatment plans based on AI's recommendations, saving valuable time and increasing the efficiency of diagnosis and treatment¹⁰.

AI technology can also discover tumor features and corresponding treatment plans through deep learning and big data analysis of tumor single-cell sequencing data. Traditional diagnostic processes are often limited by doctors' experience and knowledge, whereas AI can identify new tumor features and potential treatment methods by analyzing vast amounts of data, providing doctors with more references and options. By improving diagnostic processes, enhancing doctors' work efficiency, and discovering new tumor features and treatment methods, AI brings hope for improving patients' survival rates and quality of life.

4.2 Optimization and adjustment of treatment strategies

In the analysis of tumor single-cell sequencing data, AI not only helps to uncover more data features but also provides more personalized treatment strategies for diagnosis and treatment. In optimizing treatment strategies, AI can accurately identify the characteristics of different types of tumor cells through deep learning and pattern recognition of massive tumor single-cell sequencing data, providing clinical doctors with a more comprehensive diagnostic and therapeutic basis. Based on these features, AI can help doctors optimize treatment plans, including surgery, chemotherapy, and targeted therapy, to tailor individualized treatment plans for each patient, improving treatment efficacy. AI can also predict tumor evolution trajectories based on patients' genotypic and phenotypic data and the characteristics of tumor development, helping doctors to adjust treatment plans, thus avoiding blind treatment and unnecessary side effects.

AI can also integrate tumor single-cell sequencing data from different regions globally to establish a more comprehensive tumor database, providing clinical doctors with more extensive diagnostic references. In optimizing and adjusting treatment strategies, AI can provide more precise treatment recommendations based on the characteristics of global tumor data, achieving intelligent and globalized diagnosis and treatment. By understanding tumor development patterns and variation characteristics through AI's role in analyzing tumor single-cell sequencing data, it can provide clinical doctors with more accurate diagnostic and therapeutic information, helping doctors to be more scientific and precise when formulating treatment plans. AI offers new possibilities for personalized diagnosis and treatment of tumor patients and is expected to become an important auxiliary means in future tumor diagnosis and treatment.

5. CONCLUSIONS

AI has enormous potential and application prospects in the optimization of tumor single-cell sequencing data features

and diagnosis and treatment. Optimized data from tumor single-cell sequencing characteristics can more accurately reflect the features of tumor cells, providing a more reliable foundation for personalized treatment and precision medicine. The application of AI in tumor diagnosis and treatment can effectively improve the accuracy of diagnosis and the effect of treatment. By analyzing and learning from large amounts of data, AI can help doctors quickly determine the condition and formulate treatment plans, thereby increasing patients' survival rates and quality of life. However, there are still challenges and issues in the optimization of tumor single-cell sequencing data features and diagnosis and treatment using AI, such as data privacy protection and the interpretability of algorithms, which require further research and resolution.

ACKNOWLEDGEMENTS

This work was supported by the Fundamental Research Funds for the Central Universities of Inner Mongolia Autonomous Region, China (Grant No. GXKY22132); Natural Science Foundation of Inner Mongolia, China (Grant No. 2023LHMS03051); The central government guides local funds for scientific and technological development, China (Grant No. 2022ZY0030).

REFERENCES

- [1] Gerwin, H., Thorsten, F., Alexander, M. G., et al., "New perspectives on biology, disease progression, and therapy response of head and neck cancer gained from single cell RNA sequencing and spatial transcriptomics," *Oncology Research*, 32(1), 1-17 (2024).
- [2] Han, Y. Y., Wang, D., Peng, L. S., et al., "Single-cell sequencing: a promising approach for uncovering the mechanisms of tumor metastasis," *Journal of Hematology & Oncology*, 15, 1-19 (2022).
- [3] Xu, Y., Su, G. H., Ma, D., et al., "Technological advances in cancer immunity: from immunogenomics to single-cell analysis and artificial intelligence," *Signal Transduction and Targeted Therapy*, 6, 1-23 (2021).
- [4] Gunjan, D., Ashna, G., Tariq, M., et al., "Harnessing the potential of CAR-T cell therapy: progress, challenges, and future directions in hematological and solid tumor treatments," *Journal of Translational Medicine*, 21, 1-36 (2023).
- [5] Lei, X., et al., "Immune cells within the tumor microenvironment: Biological functions and roles in cancer immunotherapy," *Cancer Lett.*, 470, 126-133 (2020).
- [6] DeBerardinis, R. J., "Tumor microenvironment, metabolism, and immunotherapy," *N. Engl. J. Med.*, 382, 869-871 (2020).
- [7] van de Wall, S., et al., "Sialoglycans and siglecs can shape the tumor immune microenvironment," *Trends Immunol*, 41, 274-285 (2020).
- [8] Tang, T. Y., et al., "Advantages of targeting the tumor immune microenvironment over blocking immune checkpoint in cancer immunotherapy," *Signal Transduct. Tar.*, 6, 1-13 (2021).
- [9] Labani-Motlagh, A., Ashja-Mahdavi, M., Loskog, A., "The tumor microenvironment: a milieu hindering and obstructing antitumor immune responses," *Front. Immunol.*, 11, 940 (2020).
- [10] Yu, J., et al., "Liver metastasis restrains immunotherapy efficacy via macrophage mediated T cell elimination," *Nat. Med.*, 27, 152-164 (2021).