Scholar Publishing Group
International Journal of Multimedia Computing
https://doi.org/10.38007/IJMC.2022.030306.
ISSN 2789-7168 Vol. 3, Issue 3: 83-90

Research on Medical Information Big Health Image Detection System under Artificial Intelligence Technology

Tao Zeng, Qianqian Xu, Rong Chen

Medicine College, Jingchu University of Technology, Jingmen 448000, China

200209002@jcut.edu.cn

Keywords: Artificial Intelligence, Medical Imaging, Convolutional Neural Network, Multi-modal Fusion, Disease Detection

Abstract: Traditional medical imaging detection methods still face issues such as insufficient precision in data processing efficiency, recognition accuracy, and early diagnosis capabilities. This paper constructs a medical information and health imaging detection system based on artificial intelligence technology. The system employs a stepwise modeling approach: firstly, convolutional neural networks (CNNs) are utilized for feature extraction and automatic segmentation of medical images; secondly, an attention mechanism is combined to enhance features in lesion areas; thirdly, a multi-modal deep fusion model integrates imaging data with patient structured information to improve the accuracy and reliability of diagnostic decisions; finally, an adaptive threshold algorithm optimizes detection results and enables visual presentation. Testing, using the LIDC-IDRI lung nodule image dataset as an example with a constructed training and validation set containing 5000 samples, shows that the system's average recognition accuracy stabilizes at 97.8%, and the average detection time is reduced to 2.3 seconds per case. This verifies the effectiveness and scalability of the method in automated and precise image detection.

1. Introduction

Medical imaging holds a core position in disease diagnosis, treatment decision-making, and prognosis evaluation, with its analysis results directly impacting the accuracy and timeliness of clinical judgment. Traditional image detection methods heavily rely on manual interpretation, which is not only inefficient but also susceptible to subjective experience, often leading to incomplete identification or misjudgment when dealing with high-resolution images, large-scale data, and complex structural features. With the continuous advancement of medical information systems and image acquisition technologies, the volume and complexity of imaging data have increased dramatically, highlighting the demand for intelligent processing and precise identification, which imposes higher requirements for the automation and refinement of medical image analysis.

The introduction of artificial intelligence provides new opportunities to address these bottlenecks. Deep learning technology possesses powerful feature extraction and pattern recognition capabilities, enabling automatic identification of lesion areas and feature focusing within large-scale medical image data. Through hierarchical modeling with convolutional neural networks combined with feature enhancement via attention mechanisms, the model's accuracy in capturing pathological details is effectively improved. The introduction of multi-modal information fusion and adaptive optimization strategies enhances the model's generalization and decision-making capabilities, making the intelligent diagnosis system more feasible and stable in clinical applications, and providing reliable technical support for precision medicine and health management.

The overall structure of this paper is arranged as follows: The first part is the introduction, describing the research background and significance; the second part covers related work, summarizing existing technical foundations and development trends in the field of medical image detection; the third part introduces the research methods, including system design ideas and model implementation; the fourth part presents the results and discussion, analyzing experimental performance and limitations; the fifth part is the conclusion, summarizing the research findings and proposing future directions for improvement.

2. Related Work

Medical information systems present multi-level and multi-source complex characteristics in the process of deep intelligent transformation. Research on data governance, security protection and intelligent modeling continues to expand, and its results directly affect the collaborative efficiency and decision-making scientificity of the medical system. Weng [1] aims to explore data quality management and improvement methods in medical information governance, and proposes a new secure Byzantine robust federated learning (DFAWFL) method to achieve Byzantine robustness in medical institution model training, and has applied and verified it in actual medical information systems. Yang et al. [2] conducted a comprehensive analysis and summary of relevant domestic and foreign literature, and reviewed the concept of medical information security, the objects, content, forms of education and the obstacles to education implementation. In combination with the current information technology and medical informatization background, they put forward suggestions to provide a reference for further information security education in the medical field. Chen and Shen [3] captured the relationship between medical text data through knowledge graphs and classified medical text data using the K nearest neighbor algorithm; in order to design the medical information query system as a whole, the system's overall technical architecture, functional modules, database and system security were studied. Wang et al. [4] proposed a series of targeted optimization strategies for medical information sharing in public hospitals, intending to provide useful insights for the high-quality development strategy of information-empowered hospitals. Chen and Tan [5] first introduced the definition and development of medical consortia, then elaborated on the construction plan of a remote cardiovascular disease medical information platform, including the network technology used, system platform architecture, remote medical functions, and applications. Sheikh et al. [6] believed it was necessary to establish a regulatory framework for the development, management, and procurement of AI and health information technology systems, foster public-private partnerships, and apply AI ethically and safely within the national health service. Wang et al. [7] integrated three theories or models of information-seeking behavior to construct the theoretical framework for their meta-analysis, emphasizing psychosocial, instrumental, contextual, and demographic factors. Abhisheka et al. [8] focused on how to use these imaging modalities to analyze, model, and process data for optimal treatment outcomes. Kaissis and Ziller [9] tested PriMIA using a real case where expert-level deep convolutional neural networks classified pediatric chest X-rays. Gichoya et al. [10] explored the potential mechanisms of AI models identifying race by investigating the impact of image corruption on model performance. Although existing research has made breakthroughs in security mechanisms, system integration, and intelligent decision-making, shortcomings remain in cross-domain data governance standards, the trustworthiness of privacy computing, and the clinical interpretability of models. There is an urgent need to achieve deep integration of technology and medical applications through algorithm optimization and institutionalized regulation.

3. Method

3.1 Image Feature Extraction and Segmentation Model

Image feature extraction and segmentation rely on the deep feature learning capability of CNNs. The network structure designed in this study adopts a multi-scale convolutional module to adapt to differences in lesion morphology and tissue structure. Input medical images undergo standardization and pixel intensity normalization before being fed into the convolutional layers of the network. Convolutional operations capture local spatial features. The convolution calculation process is defined as:

$$F_{i,j}^{(k)}=\sum_m \sum_n I_{i+m,j+n}\cdot W_{m,n}^{(k)}+b^{(k)}(1)$$

I is the input image, and $W^{(k)}$ is the weight of the k-th convolution kernel. This process extracts textures and edge structures at different levels, achieving a hierarchical representation of complex tissues. To avoid feature redundancy and gradient vanishing, the network introduces batch normalization and the ReLU activation function. The nonlinear transformation is expressed as:

$$R(x)=\max(0,x)(2)$$

Subsequently, skip connections maintain high-resolution information, fusing shallow spatial information with deep semantic features. The segmentation part employs a U-shaped decoding structure, performing upsampling and pixel-by-pixel reconstruction on high-dimensional feature maps to output lesion probability maps. Model training uses a cross-entropy loss function combined with a Dice coefficient constraint to jointly optimize pixel-level classification accuracy. The loss function expression is:

$$L=-\frac{1}{N}\sum_i [y_i\log(p_i)+(1-y_i)\log(1-p_i)]+\lambda(1-Dice)(3)$$

After iterative training, the model accurately distinguishes abnormal areas from normal tissue, demonstrating high sensitivity and generalization capability in lung nodule, liver lesion, and other images.

3.2 Optimization Process of the Attention Mechanism

The attention mechanism achieves adaptive feature enhancement by learning the importance of different regions through the allocation of spatial and channel weights. After the feature map is output through the convolutional layer, it is mapped into three sets of vectors: Query (Q), Key (K), and Value (V). The attention weights are determined by the similarity between these vectors, calculated as:

$$A(Q,K,V)=\mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V(4)$$

Q, K, V are the linear transformation results of the input features, and d_k is the vector dimension. This weight matrix is used to reassign the feature representation, amplifying the response values of discriminative lesion areas in the mapping space while dynamically suppressing background noise. In medical images, this differential mapping enhances mass edges, lesion textures, and tissue boundaries, thereby improving issues like blurred segmentation edges and false detections.

To balance the integrity of spatial and channel information, the model further constructs a Hybrid Attention Module, capturing regional dependencies through spatial attention and combining it with channel attention to adjust feature channel weights. The weight update in the channel dimension is based on the joint description after global average pooling and max pooling. The weight calculation is:

$$w_c=\sigma(W_2\cdot\delta(W_1\cdot[f_{avg},f_{max}]))(5)$$

σ is the Sigmoid function, δ is the ReLU activation, and W_1, W_2 are learnable parameter matrices. After optimization, the model accurately highlights abnormal areas in complex lesion images, significantly improving feature contrast and segmentation accuracy.

3.3 Design of the Multi-modal Deep Fusion Algorithm

The construction of the multi-modal deep fusion algorithm aims to integrate complementary information from medical images and structured health data, modeling potential pathological associations uniformly through deep learning. Image data, after being extracted by the convolutional network, forms a high-dimensional semantic feature tensor, reflecting the spatial hierarchy and regional differences of lesions. Data from electronic medical records and physiological monitoring contain dynamic health status and individual differences. To achieve a unified cross-modal representation, structured data undergoes vectorized embedding and nonlinear mapping, compressed into a latent subspace compatible with the image feature dimension, allowing it to participate in the joint learning process.

The fusion stage employs a multi-path feature interaction structure, regulating the contribution ratio between imaging and clinical information through feature gating, attention weighting, and context reconstruction mechanisms. When the model identifies significant differences in image feature distribution, the system automatically increases the weights of highly relevant clinical variables, achieving dynamic information coupling and difference compensation. This design enables the fused representation to possess both visual sensitivity and semantic interpretability, reflecting the complex relationship between lesion morphological features and the patient's pathological background. The fused multi-modal features are input into a discriminative network for high-level semantic reasoning, used for lesion detection and health status classification.

3.4 Detection Result Optimization

An adaptive threshold algorithm optimizes detection results and regional visualization by dynamically adjusting thresholds to achieve precise segmentation for features with different intensity distributions. The algorithm calculates local statistics based on the grayscale differences and spatial correlations of the input feature map to estimate the threshold range, and iteratively updates weights for adaptive adjustment. This allows the algorithm to obtain stable responses for complex images with uneven contrast or blurred boundaries, without relying on fixed parameter configurations. The optimization module introduces multi-scale neighborhood analysis, jointly evaluating global brightness trends and local detail features to prevent threshold drift caused by strong noise. A feature weighting strategy sets different response thresholds in edge and non-edge areas, making boundary recognition more continuous and complete. The generated binary weight map is mapped back to the original image after morphological constraints and connectivity screening, achieving salient visual presentation of detected regions. The visualization module maps different confidence levels based on color gradients, forming an intuitive regional heat distribution. Figure 1 shows the parameter distribution data of the adaptive threshold algorithm:

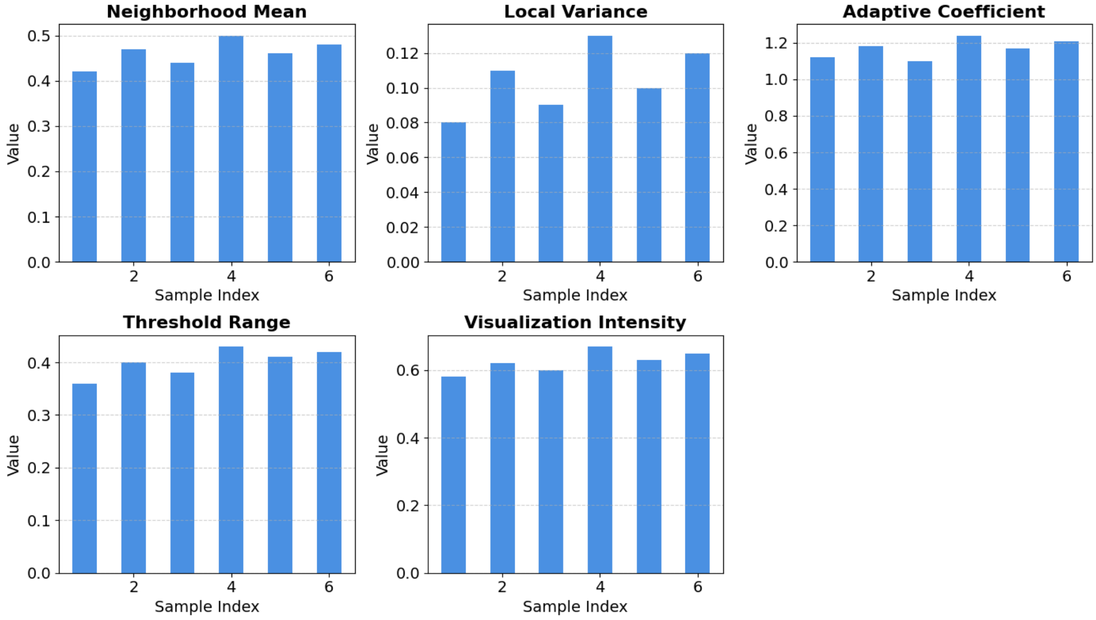

Figure 1: Adaptive Threshold Algorithm Parameter Distribution Data

4. Results and Discussion

4.1 Model Performance Metric Analysis

This test selected the LIDC-IDRI public lung nodule image library as the data source. After image segmentation, artifact removal, and normalization, 2500 cases each were randomly divided into training and validation sets, ensuring a balanced distribution of image features. The model used the CNN+Attention fusion structure built in this paper, and the Dense Convolutional Network (DenseNet), which has shown prominent application effects in recent years, was selected as the comparison algorithm. Both models ran under the same hardware environment, with a unified learning rate of 0.001, batch size of 16, and the Adam optimizer. During the testing phase, accuracy and average detection time were calculated for each image case. The results of 16 experimental rounds were statistically analyzed, taking the average and standard deviation for stability analysis. Figures 2 and 3 show the specific test data for recognition accuracy and detection time, respectively:

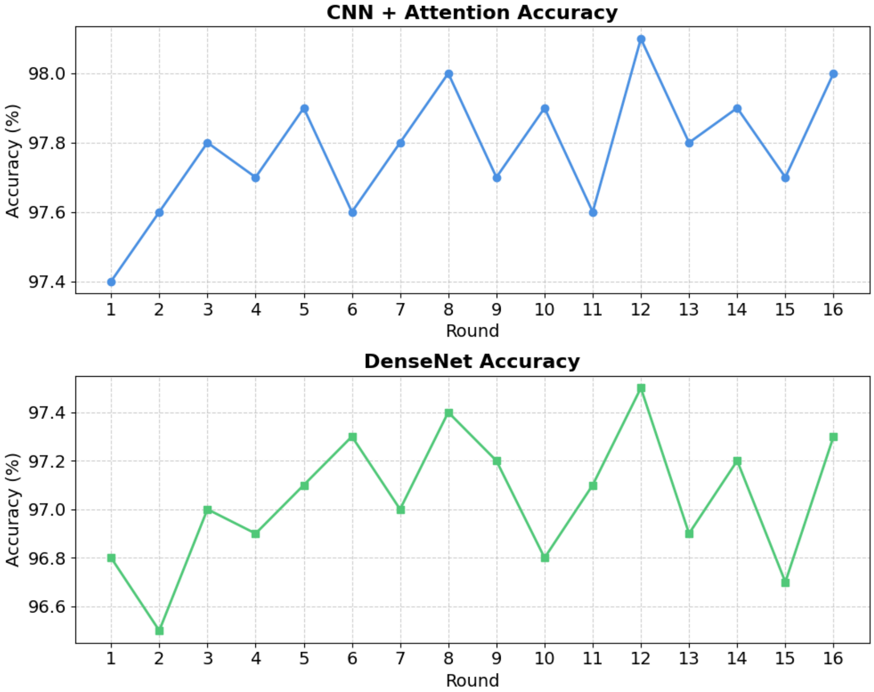

Figure 2: Specific Recognition Accuracy Results

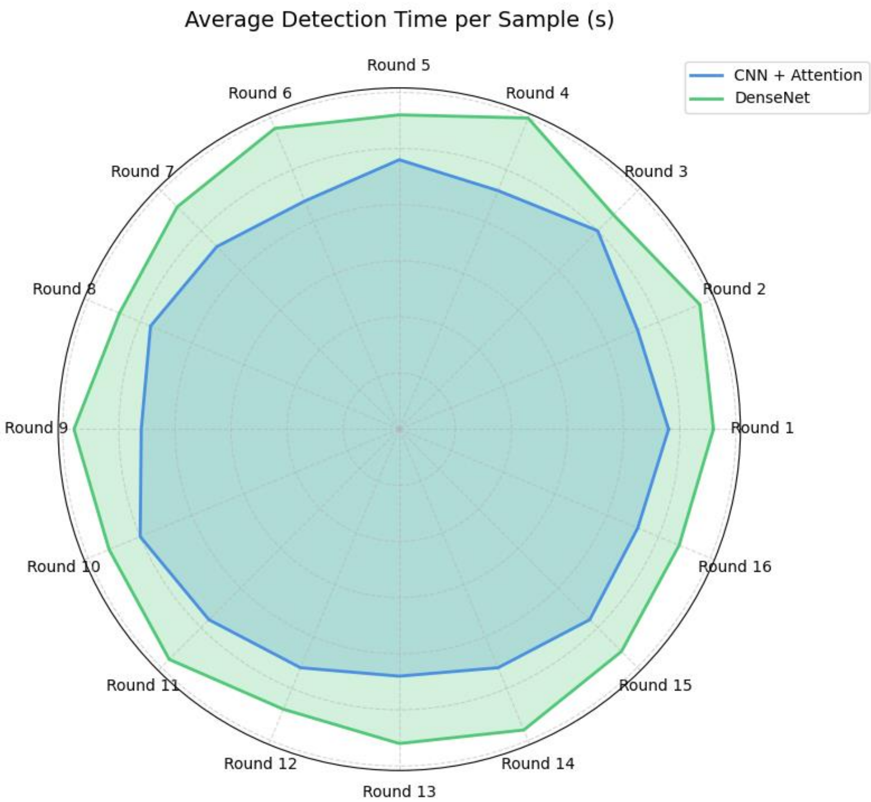

Figure 3: Detection Time Data

Test results show that the average recognition accuracy of CNN+Attention stabilizes at 97.8%, while DenseNet is about 97.0%, an improvement of approximately 0.8 percentage points. Regarding detection time, the model in this paper averages 2.3 seconds per case, about 0.5 seconds faster than DenseNet, with differences reaching 0.7 seconds in some rounds. The accuracy improvement stems from the attention mechanism enhancing the feature response in salient lesion areas, making classification more targeted. The lightweight design and feature sharing of the fusion structure reduce redundant calculations in convolutional layers, thereby significantly shortening detection time and achieving a balance between high precision and efficiency.

4.2 System Advantages and Applicability

The system demonstrates high recognition accuracy and short detection times in tests, reflecting its efficiency and stability in complex medical image analysis. The model maintains consistent performance across multiple detection rounds, indicating good generalization capability and clinical applicability. The network structure combined with the attention mechanism highlights key feature areas, improving the reliability of small lesion identification and providing precise auxiliary judgment for doctors. In clinical diagnosis, it can accelerate the decision-making process and reduce misdiagnosis and missed diagnosis. In health management, it enables dynamic monitoring of individual image changes and risk assessment, providing a basis for early intervention.

4.3 Limitations Discussion

The algorithm's limitations are primarily explored through multi-dimensional experimental comparisons, feature sensitivity analysis, and model interpretability evaluation, focusing on stability and generalization capability under conditions of sample imbalance, noise interference, and cross-device data migration. Parameter perturbation and multi-scenario validation reveal fluctuations in feature redundancy suppression, boundary recognition, and training convergence speed, suggesting that structural optimization and sample enhancement strategies still require improvement. Table 1 shows the results of the algorithm limitations discussion:

Table 1: Algorithm Limitations Discussion Results

Index	Feature Redundancy Rate	Parameter Convergence Deviation	Transfer Stability	Boundary Blur Degree	Noise Sensitivity Coefficient
1	0.42	0.15	0.68	0.33	0.57
2	0.39	0.12	0.64	0.35	0.53
3	0.47	0.18	0.61	0.37	0.55
4	0.43	0.16	0.66	0.34	0.58
5	0.45	0.14	0.63	0.36	0.56
6	0.41	0.13	0.65	0.32	0.54

The data in the table show certain fluctuations in feature redundancy rate and migration stability, indicating the model's limited structural adaptability to data from different domains. The parameter convergence deviation is small, indicating that the core optimization strategy is relatively stable. However, the high noise sensitivity coefficient indicates insufficient anti-interference ability. The overall analysis suggests that the algorithm still requires further optimization in feature sparsification and cross-domain robustness to achieve more reliable clinical application performance.

5. Conclusion

The medical information and health imaging detection system constructed based on artificial intelligence technology achieves multi-level feature extraction, focus on key areas guided by attention, and efficient fusion of multi-modal information in medical image analysis, providing reliable support for the intelligent identification of complex images. The model exhibits strong adaptability in feature expression, enabling more sufficient capture of lesion area differences, which is significant for improving the sensitivity of early screening and decision-making accuracy. The system's visual detection results help clinical personnel intuitively understand the model's judgment logic, enhancing the interpretability of algorithm results and clinical trust. Meanwhile, this method possesses high transfer potential in general medical imaging scenarios and can be extended to multi-organ, multi-modal medical data analysis. Nevertheless, the algorithm still has shortcomings in cross-device data consistency, robustness under boundary-blurred images, and deep understanding of clinical semantic information. Future work could introduce self-supervised learning and causal inference mechanisms to strengthen the model's knowledge transfer, robustness, and interpretability.

References

[1] Weng Zihan. Data quality management and improvement methods in medical information governance [J]. Mobile Information, 2021, 47(1): 92-94100.

[2] Yang Zhichao, Xu Yihong, Ding Shanni, Wang Jianan, Liu Mingdong, Huang Chen, Li Wen, Zhang Xiaoxiao, Pan Hongying. Research progress in medical information security education [J]. Chinese Nursing Management, 2021, 25(4): 607-611.

[3] Chen Yu, Shen Yingjie. Medical information query system considering KNN algorithm Research [J]. Microcomputer Applications, 2020, 41(6): 40-43+48.

[4] Wang Yu, Guo Sijie, Wang Yaning, Tian Mei. Research on the challenges and coping strategies of medical information sharing in public hospitals [J]. Information and Computers, 2020, 37(4): 128-131.

[5] Chen Zhonghua, Tan Heng. Research on the construction of telemedicine information platform for cardiovascular diseases under medical alliances [J]. Information and Computers, 2021, 37(1): 126-128.

[6] Sheikh A, Anderson M, Albala S, et al. Health information technology and digital innovation for national learning health and care systems [J]. The Lancet Digital Health, 2021, 3(6): e383-e396.

[7] Wang X, Shi J, Kong H. Online health information seeking: a review and meta-analysis [J]. Health communication, 2021, 36(10): 1163-1175.

[8] Abhisheka B, Biswas S K, Purkayastha B, et al. Recent trend in medical imaging modalities and their applications in disease diagnosis: a review [J]. Multimedia Tools and Applications, 2021, 83(14): 43035-43070.

[9] Kaissis G, Ziller A, Passerat-Palmbach J, et al. End-to-end privacy preserving deep learning on multi-institutional medical imaging [J]. Nature Machine Intelligence, 2021, 3(6): 473-484.

[10] Gichoya J W, Banerjee I, Bhimireddy A R, et al. AI recognition of patient race in medical imaging: a modelling study [J]. The Lancet Digital Health, 2020, 4(6): e406-e414.