GEORGE BROWN PRESS
JOURNALS, BOOKS & PROCEEDINGS

Journal of Computer, Signal, and System Research

Article

Research on Optimization of Automatic Medical Image Recognition System Based on Deep Learning

Minkang Zhang [1,*]

[1] Clinical Diagnostics Group (CDG) Software Team, Bio-Rad Laboratories, CA, 94547, USA
* Correspondence: Minkang Zhang, Clinical Diagnostics Group (CDG) Software Team, Bio-Rad Laboratories, CA, 94547, USA

Abstract: With the rapid development of deep learning technology, its importance to medical image recognition has been highlighted, and medical image recognition has become the core of medical development. The automatic and autonomous learning based on deep learning has greatly improved the accuracy and speed of medical image recognition, especially in the early identification of diseases and the determination of the location of diseases. However, there are still some problems such as too little data, too complicated model design and high computational cost. This paper mainly analyzes the optimization of the automatic recognition system of medical images based on deep learning, and proposes some possible methods, such as data enhancement, model structure optimization, loss function setting and transfer learning, to improve the stability and accuracy of the system.

Keywords: deep learning; medical image recognition; system optimization

Received: 24 April 2025
Revised: 28 April 2025
Accepted: 13 May 2025
Published: 22 May 2025

1. Introduction

Due to the development of artificial intelligence technology, the application of deep learning in medical image recognition system is gradually recognized, and the automated medical image recognition system can effectively ease the work intensity of doctors and improve the identification rate of diseases and speed up the recognition speed. However, in the identification process, we are faced with the processing and recognition of complex graphics, and how to improve the recognition effect, especially the performance of data processing, model design and operational efficiency, etc. Therefore, the research ideas of deep learning in the optimization of automatic medical image recognition system are discussed in depth, and specific feasible methods for improvement are given. To provide guidance for innovative development in this field and clinical practice.

2. Basic Principle of Medical Image Recognition Technology

2.1. Classification of Medical Images

Medical image mainly refers to the information obtained by various means of imaging, which is mainly used by doctors to diagnose patients, make treatment plans and evaluate them after surgery. Medical images can be divided into different categories, the most common types are radiographs, CT, magnetic resonance images, ultrasound, positron emission computed tomography, etc. X-ray imaging uses high-energy X-rays to irradiate patients and see how well they are absorbed through parts of the body. It is commonly used to diagnose fractures or lung diseases. CT technology continuously rotates the X-ray emission source to construct a hierarchical map, so CT has a higher recognition of the imaging features of brain, lung and abdominal diseases, and can display the internal images of tissues in more detail [1]. Magnetic resonance is the use of large mass magnets and high-frequency radio waves to form a clear image, with high soft tissue resolution and neuromuscular and joint anatomical resolution. Ultrasound imaging is through the continuous emission of high-frequency sound imaging, this imaging technology is often used in pregnancy fetal monitoring, heart and abdominal organs and other examinations. PET technology uses radiation imaging of specific labeled substances to study metabolic processes and mechanisms in vivo, mainly for the discovery of cancer or cardiovascular diseases and brain diseases [2].

2.2. Principle of Image Recognition Technology

Image recognition is an important branch in the field of computer vision, which mainly analyzes and processes the information in the image automatically by computer. The purpose of image recognition in medical image analysis is to find useful information features in medical images for classification, segmentation or search. In traditional methods, features and rules are generally set by people in advance to obtain image features for subsequent analysis operations, so there are some restrictions on complex images. With the introduction of deep learning, especially the application of convolutional neural network, the performance of image recognition has been greatly improved. The convolutional neural network can learn and extract the multi-level features of the image independently, and can cope with the complex graphics in the medical image better, and has the characteristics of strong generalization ability. In the field of medical image analysis, deep learning models can use a large number of labeled data for self-learning and learn to obtain disease characteristics from original images, so as to achieve automatic disease recognition, lesion segmentation, organ classification and other functions. By constantly updating the model structure and training process, deep learning has shown great potential in improving the accuracy of medical image recognition, reducing the manual intervention and misdiagnosis rate in the medical field, and has become the main tool for current medical image processing [3].

3. Application of Deep Learning in Medical Image Recognition

3.1. Automatic Classification and Recognition of Diseases

Automatic classification and recognition of diseases is mainly based on deep learning technology to automatically classify and recognize medical images. Convolutional neural networks, as deep learning, excel at automatic learning and feature extraction as well as image recognition (see Figure 1). Convolutional neural networks can acquire and extract important features from complex medical images, which can be used to distinguish and identify the site of disease.

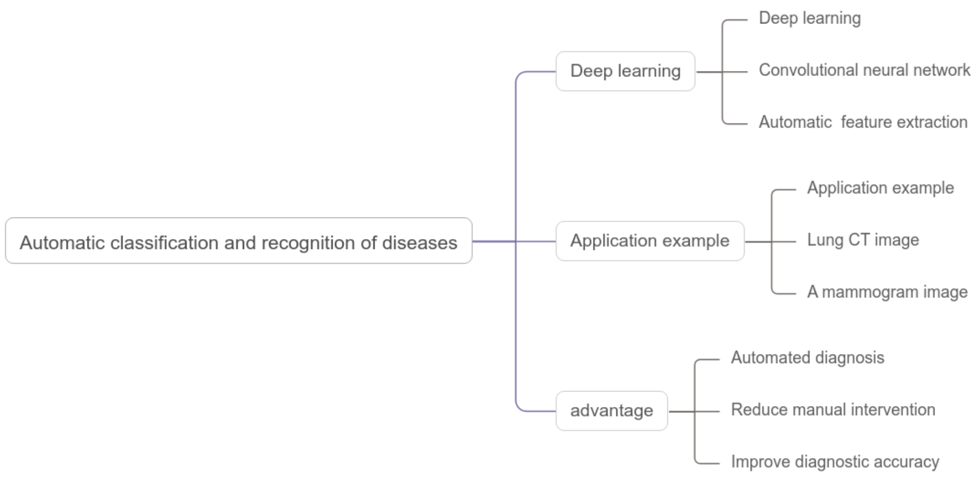

Figure 1. Automatic Classification and Recognition of Diseases.

In this way, human intervention can be reduced and diagnostic accuracy can be improved. In addition, another great advantage of deep learning is its ability to generalize and handle various forms of patient medical imaging, including the type of disease and the stage of disease. In clinical application, automatic disease recognition and classification has been widely used in many medical imaging fields. For example, in a chest CT photograph, deep learning can accurately distinguish lung nodules and identify them as benign or malignant; In mammograms, deep learning can also automatically identify signs of breast cancer, helping doctors make an initial diagnosis as soon as possible. Through deep learning technology, not only can it improve the diagnostic accuracy, but also can process a large amount of image information, reducing the work burden of doctors. Through deep learning, medical imaging research has been transformed from complex and inefficient to accurate and efficient. Automatic diagnosis system has brought strong support to the early diagnosis of diseases and promoted the development of intelligent medical treatment.

3.2. Accurate Segmentation of Lesions and Organ Regions

In disease diagnosis, accurately delineating the lesion and separating it from the surrounding organs is crucial for diagnosis and preoperative treatment. Different deep learning models have their own advantages in medical image segmentation. Through the encoder-decoder structure, U-Net is good at extracting detailed features, especially for small target segmentation. The Full Convolution Network removes the full connection layer and is suitable for pixel-level segmentation [4]. Mask R-CNN adds segmentation branch on the basis of Faster R-CNN, which can simultaneously perform target detection and pixel segmentation. DeepLabV3+ uses cavity convolution and depth-separable convolution to improve segmentation in complex backgrounds. V-Net uses 3D convolution, which is suitable for the division of lesion areas in CT, MRI and other three-dimensional images, providing more accurate diagnosis. These models can stably differentiate lesions and organ regions and provide strong support for medical image analysis. The methods and descriptions of accurate segmentation of lesions and organ regions are shown in Table 1.

Table 1. Precise Segmentation of Lesions and Organ Regions.

Method	Description
U-Net	As a mainstream image segmentation framework, U-Net adopts encoder-decoder network, which can effectively extract the details of medical images, especially for the segmentation of small targets
FCN	FullConvolutionNetwork, based on the idea of complete convolutions, is a fully connected layers that is added after the removal of conventional convolutional neural networks, and is more suitable for pixel-level image segmentation
Mask R-Convolutional neural network	Maskr-convolutional Neural Network introduces segmentation branch based on Faster-R convolutional neural network, which has the ability of object detection and pixel segmentation, so that object detection and pixel level division can be carried out at the same time
DeepLabV3+	DeepLabV3+ also uses advanced technologies such as cavity convolution and depth separable convolution to further improve the effect of image segmentation, especially for images with complex backgrounds
V-Net	V-Net uses the classic 3D convolutional network, which is good at 3D medical image segmentation, so it is more suitable for the lesion region division of CT, MRI and other 3D imaging devices, which is conducive to finding more accurate lesion location and giving more reliable disease judgment

It can be seen from the above table that deep learning has a good performance for clinical medical image segmentation, and different models have their own advantages according to different work requirements.

3.3. Detection and Localization of Lesion Target

In medical image analysis, it is a main task to find and locate the lesion target in medical image, which is very important for early diagnosis and fine treatment of disease. Object detection algorithms based on deep learning (such as YOLO, Faster-Convolutional neural networks, etc.) can automatically identify and locate diseased areas in medical images. This algorithm uses convolutional neural network to conduct in-depth analysis of images, and can identify diseased areas in artificially marked images, such as tumors, nodules, etc., and locate their coordinates in the picture. Deep learning can be used to locate lung nodules and determine the benign and malignant of lung nodules in CT images. In mammogram images, deep learning can accurately locate breast tumors and predict the benign and malignant nature of breast tumors. Compared with manual processing, deep learning improves the accuracy of detection, while shortening the detection time and reducing the probability of misjudgment. This automatic detection and positioning system can greatly reduce the burden of doctors, especially when processing a high number of medical images, it is conducive to improve the speed of doctors, shorten the diagnosis time, so as to bring faster and more accurate diagnosis and treatment experience to patients [5].

3.4. Cross-Modal Medical Imaging Assisted Diagnosis

Cross-modal medical image-assisted diagnosis is the fusion of various information of images of different modes, which makes the diagnosis of diseases more accurate and reliable. At present, the application of deep learning technology in cross-modal medical image assisted diagnosis can effectively improve the diagnosis accuracy of diseases. In clinical practice, a single modal image cannot show all pathological states, so cross-modal fusion can provide more accurate patient information for the characteristics of different images. For example, PET images show the metabolism of tumor cells, while MRI images are used to describe the size, shape and organizational structure of tumors. By fusing these two images through deep learning algorithms, questions about tumor size, location, shape and whether the tumor has spread can be obtained. Such cross-modal fusion can not only improve the diagnosis of diseases, but also improve the diagnosis of diseases. It can also provide more reliable diagnostic evidence for the development of disease treatment plans. In practical studies, cross-modal image-assisted diagnosis has become a powerful tool for precision treatment, which can help doctors understand patients more accurately and comprehensively, so as to develop more scientific and reasonable treatment methods for patients.

4. Optimization Design of Automatic Medical Image Recognition System Based on Deep Learning

4.1. Data Enhancement and Extension Methods

The method of data enhancement and extension plays an important role in medical image processing. First of all, image transformation technology includes rotation, translation and scaling and other operations, which can imitate different viewpoints, positions and scales to enhance the learning effect of the model for various deformed images. Rotation and shift can learn image features from different angles, zoom and shift can extract useful information from various positions and sizes, which is suitable for the presentation of organs and lesions in medical images. Second, data enhancement technology can enhance the data in a deeper level by means of synthetic data generation, noise addition and random clipping. Synthetic data generation can generate new training samples from existing data, especially when the amount of processing data is insufficient. Noise addition By adding random noise in the image to simulate the noise of the data in the real environment, improve the robustness of the model to noise; Random shear can make the model have strong sensitivity to respond to objects of different sizes. In short, image enhancement involves the transformation of color, light and contrast. By transforming the chroma, brightness and contrast of the image, the purpose of improving image quality and image information is achieved, and the model is assisted to obtain more accurate lesion location. Random cropping method By randomly selecting an area in the original image for cropping, the cropping operation formula can be expressed as:

$$I' = I[x{:}\, x + w, y{:}\, y + h] \quad (1)$$

(x, y) is the top-left coordinate of the crop area, w, h are the width and height of the crop area, I' is the cropped image. These mathematical formulas represent common data enhancement methods that transform the original image through different operations to generate new training samples and improve the generalization ability of deep learning models.

4.2. Model Structure Optimization and Improvement

The structural improvement and optimization of deep learning model can be regarded as one of the important bases to improve the performance of automatic medical image recognition system. Although traditional convolutional neural networks (convolutional neural networks) have good performance in image classification tasks, they are often prone to certain limitations in the face of the complexity and complex details of medical images. Therefore, it is necessary to design a more efficient network structure. For example, ResNet solves the problem of gradient disappearance in deep network training by using residual connection, which can greatly enhance the training effect. The commonly used network structure is also U-Net, which is widely used in medical image segmentation tasks. Due to the structure of encoding and decoding blocks, it can effectively solve the details and important information in medical images. In addition, the network structure based on attention is getting more and more attention, which can make the network adjust the weight of each part of the feature, make the network pay more attention to some important parts, and improve the accuracy of model identification. Perfecting and optimizing the network structure of deep learning can help enhance the understanding of the network to medical pictures, and thus improve the performance of the automatic identification system.

4.3. Loss Function Design and Optimization

In the process of training a deep learning model, the loss function is an important factor determining the learning ability of the model. Traditional cross-entropy loss and mean square error are often used in classification and regression. However, the medical image recognition task may have more complex problems, and an accurate loss function needs to be designed for the task. Taking the loss function in the medical image slicing task as an example, the Dice coefficient loss can characterize the similarity between the slice result and the real label, especially in the case of data imbalance, it can effectively avoid the problem of ignoring small targets. In the case of the fusion of each task loss function, the classification, positioning and slicing tasks can be optimized at the same time to improve the learning effect of the multi-task model. The optimization of the loss function not only optimizes the accuracy of the model, but also contributes to the stability of the model learning, reduces the risk of overfitting, and further improves the efficiency and accuracy of the automatic medical image recognition system. Dice coefficient loss is commonly used in image segmentation tasks to evaluate the performance of the model by calculating the overlap between the predicted segmentation and the actual segmentation region. The Dice coefficient loss function is as follows:

$$L_{Dice} = 1 - \frac{2\sum_{i=1}^{N} p_i y_i + \epsilon}{\sum_{i=1}^{N} p_i + \sum_{i=1}^{N} y_i + \epsilon} \quad (2)$$

Among, p_i is the pixel value of the image predicted by the model, y_i is the pixel value of the actual annotated image, N is the total number of pixels in the image, ϵ is a small constant that prevents division by zero errors. By optimizing the Dice coefficient loss, the model can better focus on the key areas in the image, improve the segmentation accuracy, and play an important role in the automatic diagnosis and treatment planning of medical images.

4.4. Model Integration and Transfer Learning

The performance of deep learning medical image recognition system can be improved by using two important technologies: model integration technology and transfer learning, the former can reduce the single model bias and improve its performance; The latter can make up for the lack of labels by using the model that has been trained in advance. After the model is trained on large datasets in advance, it can be applied in the field of medical imaging to realize the correct correction of the model, which can greatly improve the learning speed and accuracy of the model. Especially when the sample size is small, it has a great impact on the efficiency of the model. Combining the above techniques and effects can form a more powerful medical image adaptive recognition system, which can help doctors better diagnose diseases and formulate corresponding treatment methods.

5. Conclusion

By means of data enhancement, model structure optimization, loss function selection, model integration and model transfer, several effective methods to improve system performance are proposed. With the continuous development of deep learning technology, its use in the early stage of disease identification, location of the lesion area, the development of disease treatment plans, etc., is becoming more and more mature. However, there are still some problems that need to be solved, such as how to improve data quality, improve model interpretability, and how to further enhance computing resources. When more new technologies are used and larger data is accumulated, medical image recognition systems based on deep learning will have a greater impact on clinical applications, which is conducive to the development of intelligent medicine.

References

1. L. Shen, Z. Shao, Y. Yu, and X. Chen, "Hybrid approach combining modified gravity model and deep learning for short-term forecasting of metro transit passenger flows," *Transp. Res. Rec.*, vol. 2675, no. 1, pp. 25–38, 2021, doi: 10.1177/0361198120968823.
2. S. Erdem and C. Thompson, "Prioritising health service innovation investments using public preferences: a discrete choice experiment," *BMC Health Serv. Res.*, vol. 14, p. 360, 2014, doi: 10.1186/1472-6963-14-360.
3. U. Schmitt, "Validating and documenting a new knowledge management system philosophy: a case based on the ISO 30401:2018-KMS standard," *Knowl. Manag. Res. Pract.*, vol. 20, no. 6, pp. 960–974, 2022, doi: 10.1080/14778238.2022.2064349.
4. M. Wang, R. Wang, J. Liu, W. Ju, Q. Zhou, G. Zhang, and M. Wang, "Operation optimization for park with integrated energy system based on integrated demand response," *Energy Rep.*, vol. 8, pp. 249–259, 2022, doi: 10.1016/j.egyr.2022.05.060.
5. Z. Wang, W. Zhou, I. M. Jiskani, H. Luo, Z. Ao, and E. M. Mvula, "Annual dust pollution characteristics and its prevention and control for environmental protection in surface mines," *Sci. Total Environ.*, vol. 825, p. 153949, 2022, doi: 10.1016/j.scitotenv.2022.153949.