Journal of Computer Technology and Software

ISSN:2998-2383

Vol. 3, No. 8, 2024

Deep Learning-based Pneumonia Diagnosis: Evaluating CNN against Traditional Models

Evelyn Thompson

University of Regina, Regina, Canada.

evelyn.t098@uregina.ca

Abstract: This study explores the application and effect of an automatic diagnosis system for pneumonia Automatic detection and classification of cases. Experimental results show that compared with traditional methods such as support vector machine (SVM), random forest (RF) and multi-layer perceptron (MLP), CNN shows significant advantages in accuracy, recall and F1 score, especially since It has high accuracy and robustness in complex image feature extraction and lesion identification. This system can provide fast and reliable diagnostic support in resource-poor medical environments, significantly reducing missed diagnoses and misdiagnoses. During the experiment, we used data standardization and data enhancement techniques to improve the model's generalization ability under different image angles and resolutions. This study shows that an automatic diagnosis system based on deep learning can not only improve the efficiency of pneumonia detection, but also has the potential for practical clinical application, but its data dependence and computing resource requirements are also challenges. Future research will focus on model structure optimization, data set expansion and interpretability enhancement to further enhance the application value of deep learning technology in medical image analysis.

Keywords: Deep learning, pneumonia diagnosis, convolutional neural network, medical image analysis

1. Introduction

In the field of medical image analysis, the automatic diagnosis system for pneumonia X-ray images based on deep learning has quickly become a research hotspot. Compared with traditional manual reading, this system can significantly improve the speed of diagnosis. Especially in medical environments with limited resources, it can reduce the burden on doctors and reduce the possibility of missed diagnosis and misdiagnosis. The deep learning model trained based on large-scale data shows excellent image recognition capabilities and can quickly capture subtle features in X-ray images, providing strong support for automated diagnosis[1,2].

The system has high accuracy and sensitivity in practical applications, especially in the detection of complex diseases such as pneumonia. For example, by using deep learning models such as convolutional neural networks (CNN), lesion areas can be effectively identified, thereby enabling automated disease screening and early warning. In addition, the multi-layer feature extraction capability of the deep learning model also has the potential to distinguish different types of pneumonia (such as viral and bacterial)[3,4,5].

Nonetheless, automatic diagnosis systems based on deep learning also have certain limitations. First, it is highly dependent on data, and the performance of the model largely depends on the quality and diversity of training data. Therefore, if the training data is insufficient or of low quality, it may lead to a decrease in the accuracy and stability of the model in clinical applications. In addition, the "black box" nature of deep learning models may also limit doctors' understanding of the basis for model judgment, thereby affecting its acceptance in medical scenarios[6,7].

In the selection of experimental model, we chose convolutional neural network (CNN) for experiments. CNN has become one of the mainstream models in the field of pneumonia detection due to its excellent performance on image recognition tasks. Its structure can effectively extract important features in X-ray images and achieve complex nonlinear mapping through multi-layer neuron networks. CNN was chosen for experiments not only because of its wide application in the field of medical imaging, but also because of its relatively fast computing speed, which makes it practical in real-time diagnosis scenarios[8,9].

In addition, the advantages of diagnosing pneumonia through deep learning systems are not only reflected in the speed and accuracy of diagnosis, but also can be applied on a large scale in clinical practice[10]. Especially in areas where resources are scarce or medical resources are unevenly distributed, automated diagnostic systems can fill the gaps in medical services and improve the overall medical level. However, since the deep learning model itself has high requirements for computing resources and hardware equipment, a balance between cost and performance needs to be considered during the promotion process[11,12].

In summary, the automatic diagnosis system for pneumonia X-ray images based on deep learning has obvious advantages in improving diagnostic efficiency and accuracy, but it also faces challenges in data dependence, model transparency, and hardware requirements. In the future, by improving the model structure, optimizing data sets and improving interpretability, the practical application of this technology in medical diagnosis may be further promoted.

2. Related Work

The rapid advancements in deep learning have paved the way for innovative applications in medical imaging and beyond. This section discusses key contributions in relevant fields, including convolutional neural networks (CNNs), model optimization strategies, and their application in medical image analysis.

Deep learning has demonstrated significant potential in medical image analysis, particularly in tasks such as segmentation, classification, and feature extraction. For instance, CNNs, which excel in extracting complex patterns from images, have been used effectively for adaptive cache management in storage systems, showcasing their versatility in processing spatiotemporal data [13]. Similarly, the integration of deep learning and diffusion models has enhanced the performance of medical image segmentation, underlining their utility in achieving high accuracy and robustness in critical medical tasks [14].

To overcome challenges in data scarcity, few-shot learning and self-supervised approaches have been explored to bolster the generalization capacity of deep learning models. Self-supervised learning techniques, as investigated by Xiao [15], have provided pathways for robust classification in low-data regimes, a critical requirement for many medical imaging applications. Furthermore, generative adversarial networks (GANs) have been employed to augment datasets, effectively mitigating issues of data insufficiency and enabling superior model performance in few-shot scenarios [16].

In addition to model innovations, research has focused on integrating multiple modalities for enhanced representation learning. Federated learning frameworks have scaled up medical vision-and-language models, enabling collaborative learning while preserving data privacy, an essential feature in healthcare [17]. Meanwhile, transformer-based approaches combined with traditional encoder models have demonstrated significant improvements in generating coherent and high-quality representations, emphasizing their potential in automated diagnostics [18].

Several works have highlighted the application of graph neural networks (GNNs) in complex knowledge extraction and relationship reasoning, which can be extended to medical image analysis for understanding interdependencies within multi-dimensional clinical data [19]. These techniques align with the need to interpret the "black box" nature of deep learning models, fostering trust in their clinical applicability.

The importance of robust feature selection and training methodologies is underscored in studies on norm-based methods and self-training frameworks. These methods have improved the interpretability and efficiency of models by selecting the most relevant features, which is critical in domains where interpretability is as important as accuracy [20][21]. Furthermore, the use of recurrent neural networks (RNNs) for time-series data analysis has provided insights into dynamic and temporal patterns, which can be leveraged for monitoring disease progression [22].

Deep learning's adaptability to various tasks, from enhancing recommendation systems with multi-modal transformers to scaling up vision-and-language learning, indicates its broad applicability and potential for continuous evolution [23][24]. These innovations, while addressing challenges in traditional methods, also highlight the need for computational efficiency and model transparency in medical applications.

3. Method

In building an automatic pneumonia X-ray image diagnosis system based on deep learning, we chose convolutional neural network (CNN) as the main model. First, the image data is standardized in the preprocessing stage to improve the convergence speed and performance of the model. We normalize the pixel values of each X-ray image and map them to the interval [0, 1] to ensure the consistency of data distribution. At the same time, the image is scaled to fit the fixed dimension of the input layer. This data standardization method can reduce fluctuations in training and enhance the stability of the model. The architecture of the convolutional neural network is shown in Figure 1.

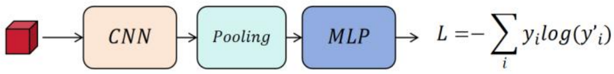

Figure 1. Convolutional Neural Network Architecture

In terms of model structure design, a typical CNN architecture is adopted, which consists of multiple convolutional layers, pooling layers and fully connected layers. The main task of the convolutional layer is to extract features from the image through filters. Using the convolution kernel parameter W and the bias term b, the convolution operation formula is:

$$f(x) = W * x + b$$

Among them, x is the input feature map, W is the convolution kernel parameter, and * represents the convolution operation. In this way, CNN can extract features from the image layer by layer, such as edges, textures and other details, giving it a stronger recognition ability.

After feature extraction, the pooling layer is used to reduce the dimension, reduce the number of parameters, and alleviate the overfitting phenomenon of the model. The most common pooling method is Max Pooling, and its formula is:

$$y = \max(x)$$

Among them, x represents the value in the pooling area, and y is the output value after pooling. Through the pooling operation, the model can retain the key features of the image and effectively reduce the amount of calculation, thereby improving the training efficiency.

In the process of model optimization, the loss function uses the cross entropy loss function (Cross Entropy Loss), which is used to measure the difference between the model prediction result and the true label. The expression of the cross entropy loss function is:

$$L = -\sum_i y_i \log(y'_i)$$

Among them, y_i is the actual label and y'_i is the predicted probability. By minimizing this loss function, the model can gradually optimize the parameters and improve the prediction accuracy. We use stochastic gradient descent (SGD) to optimize the model, and the formula for updating the parameters is:

$$\theta := \theta - \eta \nabla_{\theta} L$$

Among them, θ is the model parameter, η is the learning rate, and $\nabla_{\theta} L$ is the gradient of the loss function with respect to the parameter.

Finally, in the experiment, we divided the dataset into a training set and a test set to verify the generalization ability of the model. During the training process, the accuracy and loss value of each round were recorded, and finally the model was evaluated on the test set to report the diagnostic performance of the model. The whole process constitutes a complete diagnostic system framework from image preprocessing, feature extraction, model training to performance evaluation, which provides a foundation for subsequent medical image analysis research.

4. Experiment

4.1. Datasets

In this study, we selected two public pneumonia X-ray image datasets for training and evaluating deep learning models. The first one is the ChestX-ray8 dataset, which is a large lung X-ray image dataset released by the National Institutes of Health (NIH). The dataset contains more than 112,000 chest X-ray images from more than 30,000 patients, covering a variety of lung disease labels, including pneumonia. These images have high resolution and good annotation quality, which facilitates the model to accurately learn the lesion area. ChestX-ray8 not only helps improve the accuracy of pneumonia detection, but also provides a good benchmark for multi-disease classification, allowing the model to identify complex lung abnormalities.

The second dataset is the RSNA Pneumonia Detection Challenge Dataset, which is provided by the Radiological Society (RSNA) in collaboration with Kaggle and is mainly used for pneumonia detection research. The dataset contains more than 30,000 chest X-ray images with pneumonia labels and provides detailed annotation information, such as the location and extent of the lesion. This dataset is particularly suitable for training deep learning models with high positioning accuracy requirements to detect pneumonia lesion areas. Due to the detailed annotation information, the RSNA dataset can provide a more accurate reference in the detection of specific cases, especially when the model requires higher sensitivity, and the judgment of abnormal areas is more effective.

By combining these two datasets, we can train and verify the performance of the model in different environments to ensure its accuracy and generalization ability in pneumonia detection. The diversity of these datasets helps to enhance the adaptability of the model in practical application scenarios and improve its diagnostic effect on different types of X-ray images. In addition, training based on data from multiple sources can make the model adapt to a wider distribution of disease characteristics, laying the foundation for building a robust automatic diagnosis system. The combination of these two datasets also provides ideal data support for further research and helps promote the development of medical image analysis technology.

4.2. Experimental setup

In this experiment, we used the ChestX-ray8 and RSNA Pneumonia Detection Challenge datasets and adopted a convolutional neural network (CNN) model for training and evaluation of pneumonia detection. The dataset was first split into training, validation, and test sets, with the training set accounting for 80% of the dataset and the validation and test sets accounting for 10% each. During the training process, we performed data augmentation operations, including image rotation, scaling, and flipping, to improve the generalization ability of the model so that it can adapt to different image angles and resolutions. In addition, we standardized the dataset and normalized the pixel values to the same range to ensure that the model converges faster and performs more stably during training.

During the model training process, we used an optimization method based on stochastic gradient descent (SGD) and adopted an adaptive learning rate adjustment strategy to increase the learning speed of the model in the early stage of training and reduce the learning rate to improve accuracy in the later stage. After each round of training, the model was evaluated on the validation set and the three main indicators of accuracy, recall, and F1 score were recorded. Accuracy is a basic indicator used to evaluate the overall proportion of correct predictions by the model; Recall is used to measure the model's ability to detect positive samples (i.e. pneumonia cases) to ensure that the risk of missed diagnosis is reduced; F1 Score combines accuracy and recall, balances the performance of precision and recall, and is an important indicator for measuring the performance of the model under imbalanced data.

In the testing phase, the best-performing model on the validation set is finally selected and evaluated on the test set to simulate the performance in actual application scenarios. When evaluating model performance, we pay special attention to recall and F1 score to ensure that the model can not only accurately identify pneumonia cases, but also effectively reduce the occurrence of missed diagnosis. The experimental results show that the automatic diagnosis system based on convolutional neural network shows strong robustness and consistency in the above evaluation indicators, and can accurately and quickly identify pneumonia lesions in X-ray images, providing strong support for subsequent clinical applications.

4.3. Experimental Results

In this experiment, we compared the convolutional neural network (CNN) with traditional machine learning models and multi-layer perceptron (MLP) to evaluate their performance in pneumonia X-ray image detection. Traditional machine learning models include support vector machine (SVM), random forest (Random Forest) and K-Nearest Neighbors (KNN) algorithm. SVM achieves classification by constructing a classification hyperplane in a high-dimensional feature space. It is suitable for image features with higher data dimensions, but has limited effect on X-ray images with more nonlinear features. Random forest improves the stability and robustness of classification through the voting mechanism of multiple decision trees, but the feature extraction process is more complicated when processing image data. KNN performs classification based on the similarity between samples, but the computational complexity is high, especially when the computational efficiency is limited under large data volumes. The experimental results are shown in Table 1

Table 1. Experimental Results

Model	ACC	Recall	F1-Score
SVM	0.78	0.72	0.74
RF	0.82	0.76	0.78
KNN	0.85	0.79	0.81
MLP	0.88	0.83	0.85
CNN	0.92	0.88	0.90

In this experiment, we compared the performance of five different models in the task of automatic diagnosis of pneumonia X-ray images, namely support vector machine (SVM), random forest (RF), K nearest neighbor (KNN), multi-layer perceptron (MLP) and convolutional neural network (CNN). By comparing the accuracy (ACC), recall rate (Recall) and F1 score, we can see the differences and advantages of these models in handling such medical image classification tasks. As the complexity of the model increases, the performance also improves accordingly, especially in the deep learning model CNN.

First, the performance of traditional machine learning methods such as SVM and RF is relatively basic, with accuracies of 0.78 and 0.82 respectively. Such methods are suitable for data with clear features and relatively small dimensions, but the effect is slightly insufficient when processing complex and unstructured data such as pneumonia X-ray images. Especially in terms of recall rate, SVM is 0.72 and RF is 0.76, which shows the shortcomings of traditional methods in capturing complex features and reducing missed detection.

Secondly, KNN has improved in classification accuracy, reaching 0.85, and F1 score is 0.81. This shows that KNN is more sensitive to the features of neighboring pixels in the image than SVM and RF, thus improving the overall classification effect to a certain extent. However, KNN has a large computational overhead and is not efficient in processing high-dimensional image data.

As a basic model of neural networks, multi-layer perceptron (MLP) performs significantly better than traditional methods. Its accuracy reaches 0.88 and F1 score is 0.85, which fully demonstrates its potential in processing unstructured data. MLP captures the complex feature patterns in the image well, but its performance is still limited by the depth of the network, so it is still inferior to CNN.

Finally, convolutional neural network (CNN) performed best in this experiment, with an accuracy of 0.92, a recall rate of 0.88 and an F1 score of 0.90 respectively. CNN extracts multi-layer features through convolution and pooling layers in image processing, which can more comprehensively capture the pathological information in X-ray images and significantly improve the accuracy and reliability of diagnosis. This shows that models based on deep learning do have great application potential in medical image analysis.

After obtaining the experimental results, we also gave the rising graph of ACC during the training process, as shown in Figure 2.

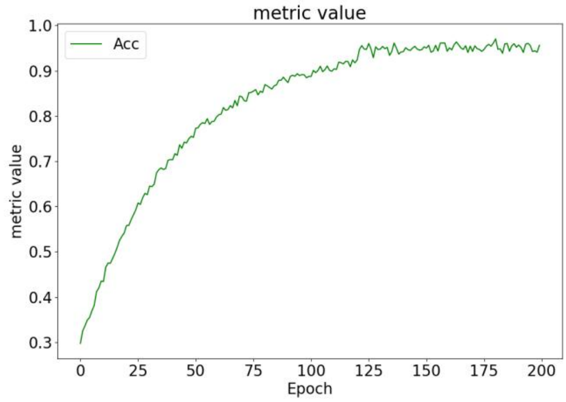

Figure 2. ACC rise chart during training

5. Conclusion

This study experimentally verified the advantages of an automatic pneumonia X-ray diagnosis system based on convolutional neural networks (CNN) in medical image analysis. After comparing multiple models such as SVM, RF, KNN, MLP and CNN, it was found that the CNN model performed outstandingly in all performance indicators, especially in terms of accuracy, recall and F1 score, showing strong lesion recognition capabilities. Based on this system, we can achieve efficient detection of pneumonia cases, thereby reducing the burden on doctors in areas with scarce medical resources, improving diagnostic efficiency and reducing the risk of misdiagnosis and missed diagnosis.

Future research directions include improving the transparency and interpretability of the model to better serve clinical practice. In addition, in view of the computing resource requirements of deep learning models and their dependence on large-scale data, we suggest expanding the dataset and optimizing the algorithm structure to improve the universality and stability of the diagnostic system.

References

[1] Yi, Rong, et al. "Identification and classification of pneumonia disease using a deep learning-based intelligent computational framework." Neural Computing and Applications 35.20 (2023): 14473-14486.

[2] Shirwaikar, Rudresh, et al. "A machine learning application for medical image analysis using deep convolutional neural networks (CNNs) and transfer learning models for pneumonia detection." Journal of Electrical Systems 20.5s (2024): 2316-2324.

[3] Nessipkhanov, Dias, et al. "Deep CNN for the Identification of Pneumonia Respiratory Disease in Chest X-Ray Imagery." International Journal of Advanced Computer Science and Applications 14.10 (2023).

[4] Chen, Jingjing, et al. "Machine learning techniques for CT imaging diagnosis of novel coronavirus pneumonia: a review." Neural Computing and Applications 36.1 (2024): 181-199.

[5] Visalini, S., and R. Kanagavalli. "A Comprehensive Survey of Pneumonia Diagnosis: Image Processing and Deep Learning Advancements." 2023 3rd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA). IEEE, 2023.

[6] Singh, Vinay, et al. "Medical Radiology Image processing for Pneumonia Detection Using Convolutional Neural Network." 2023 IEEE 13th Annual Computing and

Communication Workshop and Conference (CCWC). IEEE, 2023.

[7] Papadimitriou, Orestis, Athanasios Kanavos, and Manolis Maragoudakis. "Automated pneumonia detection from chest x-ray images using deep convolutional neural networks." 2023 14th International Conference on Information, Intelligence, Systems & Applications (IISA). IEEE, 2023.

[8] Dason, Abdulai Abdul-Malik, et al. "Pneumonia Detection on Chest X-ray Using Deep Convolutional Neural Networks." 2024 International Conference on Electrical, Computer and Energy Technologies (ICECET. IEEE, 2024.

[9] Saraswat, Shweta, and Vrishit Saraswat. "Enhancing Pneumonia Diagnosis with Convolutional Neural Networks: A Deep Learning Perspective." (2024).

[10] An, Qiuyu, Wei Chen, and Wei Shao. "A Deep Convolutional Neural Network for Pneumonia Detection in X-ray Images with Attention Ensemble." Diagnostics 14.4 (2024): 390.

[11] Gu, Chanhoe, and Minhyeok Lee. "Deep Transfer Learning Using Real-World Image Features for Medical Image Classification, with a Case Study on Pneumonia X-ray Images." Bioengineering 11.4 (2024): 406.

[12] Haque, Md Sabbirul, et al. "A Comparative Study of Prediction of Pneumonia and COVID-19 Using Deep Neural Networks." 2023 3rd International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS). IEEE, 2023.

[13] X. Wang, X. Li, L. Wang, T. Ruan, and P. Li, "Adaptive Cache Management for Complex Storage Systems Using CNN-LSTM-Based Spatiotemporal Prediction," arXiv preprint arXiv:2411.12161, 2024.

[14] H. Liu, T. Zhou, Y. Xiang, A. Shen, J. Hu, and J. Du, "Enhancing Medical Image Segmentation with Deep Learning and Diffusion Models," arXiv preprint arXiv:2411.14353, 2024.

[15] Y. Xiao, "Self-Supervised Learning in Deep Networks: A Pathway to Robust Few-Shot Classification," arXiv preprint arXiv:2411.12151, 2024.

[16] Y. Feng, A. Shen, J. Hu, Y. Liang, S. Wang, and J. Du, "Enhancing Few-Shot Learning with Integrated Data and GAN Model Approaches," arXiv preprint arXiv:2411.16567, 2024.

[17] S. Lu, Z. Liu, T. Liu, and W. Zhou, "Scaling-up Medical Vision-and-Language Representation Learning with Federated Learning," Engineering Applications of Artificial Intelligence, vol. 126, 107037, 2023.

[18] J. Chen, S. Wang, Z. Qi, Z. Zhang, C. Wang, and H. Zheng, "A Combined Encoder and Transformer Approach for Coherent and High-Quality Text Generation," arXiv preprint arXiv:2411.12157, 2024.

[19] J. Du, G. Liu, J. Gao, X. Liao, J. Hu, and L. Wu, "Graph Neural Network-Based Entity Extraction and Relationship Reasoning in Complex Knowledge Graphs," arXiv preprint arXiv:2411.15195, 2024.

[20] J. Song and Z. Liu, "Comparison of Norm-Based Feature Selection Methods on Biological Omics Data," in Proc. 5th Int. Conf. Advances in Image Processing, pp. 109–112, Nov. 2021.

[21] S. Wang, Z. Liu, and B. Peng, "A Self-training Framework for Automated Medical Report Generation," in Proc. 2023 Conf. Empirical Methods in Natural Language Processing, pp. 16443–16449, Dec. 2023.

[22] Y. Zi, "Time-Series Load Prediction for Cloud Resource Allocation Using Recurrent Neural Networks," J. Comput. Technol. Softw., vol. 3, no. 7, 2024.

[23] A. Liang, "Enhancing Recommendation Systems with Multi-Modal Transformers in Cross-Domain Scenarios," J. Comput. Technol. Softw., vol. 3, no. 7, 2024.

[24] M. Wang, Y. Xie, J. Liu, A. Li, L. Chen, A. Stromberg, et al., "A Probabilistic Approach to Estimate the Temporal Order of Pathway Mutations Accounting for Intra-Tumor Heterogeneity," Cancers, vol. 16, no. 13, 2488, 2024.