Journal of Computer Technology and Software

ISSN:2998-2383

Vol. 4, No. 2, 2025

Research on Medical Image Diagnosis Models Based on Convolutional Neural Networks

Marina Caldwell

University of Houston-Clear Lake, Houston, USA

mcaldwell@uhcl.edu

Abstract: This study explored the application of convolutional neural network (CNN) based medical image classification and compared the performance of five common deep learning models, including CNN, VGG16, ResNet-50, Inception-v3 and MobileNet. Through experiments on public medical image datasets, the classification accuracy, recall and F1 score of each model were evaluated. The experimental results show that VGG16 performs best in all evaluation indicators, with an accuracy of 90.2%, a recall of 88.5% and an F1 score of 89.3%. Although ResNet-50 also has high performance, it is slightly inferior to VGG16 in accuracy and recall. Inception-v3 and MobileNet perform relatively poorly in processing medical images, with lower accuracy and recall than the former two. CNN, as a baseline model, performs the worst in various indicators, showing its limitations in medical image analysis. Through comparative experiments, this study provides a theoretical basis for selecting appropriate deep learning models for medical image classification and provides guidance for the future application of deep learning in medical image diagnosis.

Keywords: Medical image classification, convolutional neural network, VGG16, ResNet-50, deep learning

1. Introduction

With the rapid development of artificial intelligence technology, the application of deep learning in medical image analysis has made significant progress. As an important basis for clinical diagnosis and treatment, medical imaging plays a vital role in the early detection and accurate judgment of diseases. However, due to the high complexity and high dimensionality of medical images, traditional manual analysis methods are often limited by time and accuracy and are difficult to meet the needs of modern medical diagnosis. Therefore, how to use advanced computing technology to improve the efficiency and accuracy of medical image analysis has become an urgent problem in the medical field. As a powerful deep learning tool, convolutional neural network (CNN) has been widely used in image classification, target detection and other fields, especially in medical image analysis, showing its strong potential[1].

Convolutional neural networks can automatically extract feature information from images through structures such as multi-layer convolution, pooling and fully connected layers, greatly reducing the complexity of traditional manual feature extraction. In medical image classification tasks, CNN can automatically learn important diagnostic features from original images, avoiding the tedious process of manually designing features. Therefore, the application of CNN in medical image analysis has received more and more attention, especially in tasks such as disease classification, tumor detection, and organ segmentation, where it has demonstrated its superior performance[2]. As a classic convolutional neural network model, VGG16 has become one of the important tools in the field of medical image classification with its deep network structure and excellent image feature extraction capabilities[3].

The structural characteristics of the VGG16 model enable it to have strong expressive capabilities in medical image classification tasks[4]. The model consists of 16 layers, including 13 convolutional layers and 3 fully connected layers, with deeper network depth and higher nonlinear fitting capabilities. The core advantage of VGG16 lies in its simple and efficient structural design. It uses a small-sized 3×3 convolution kernel for feature extraction, which makes the model have strong feature learning capabilities. Compared with other deep network models, VGG16 has better performance in images. Performance in classification tasks is very stable. Through reasonable network design, VGG16 can extract rich local and global features in medical images, providing strong support for subsequent classification tasks[5].

Medical imaging data often suffers from high dimensions and difficulty in labeling. The powerful feature extraction capabilities of VGG16 can effectively alleviate this problem[6]. Through training, VGG16 can automatically learn key diagnostic information from a large amount of medical image data, greatly improving the accuracy of image classification. In specific applications, VGG16 can be used to classify various medical images, including CT, MRI, X-ray and other image types. Compared with traditional methods, VGG16 can not only achieve better results in classification accuracy, but also significantly reduce the workload of manual annotation and improve the work efficiency of clinicians.

In the application process of medical image classification, the quality and quantity of the data set have a crucial impact on the training effect of the model. In order to solve the problems of difficulty in labeling medical imaging data and insufficient data, data enhancement technology and transfer learning methods have been widely used in recent years. Data enhancement can effectively expand the size of the data set and improve the generalization ability of the model by performing operations such as rotation, translation, scaling, and flipping on the original image. Transfer learning pre-trains VGG16 on large-scale image data sets and then applies it to medical image classification tasks, thereby greatly reducing training time and dependence on the amount of data. With the support of these technologies, the VGG16 model can more accurately adapt to medical image classification tasks, improving the model's performance in practical applications[7].

Although VGG16 has shown strong capabilities in medical image classification, there are still some challenges and room for improvement in practical applications. First, the computational complexity of the VGG16 model is high, especially when processing large-scale medical images, and the training and inference processes may take a long time. Secondly, due to the diversity and complexity of medical imaging data, VGG16 may not be able to fully capture all diagnostic information, causing the model's robustness and accuracy to be affected to a certain extent. Therefore, in response to these problems, researchers are exploring ways to further improve the performance of VGG16 in medical image classification by optimizing the network structure, introducing attention mechanisms, and combining multi-scale information.

In general, medical image diagnosis models based on convolutional neural networks, especially the use of VGG16 for medical image classification, have become an important tool for modern medical image analysis. Through effective feature extraction and model optimization, VGG16 can provide efficient and accurate solutions in medical image classification tasks, providing strong support for early diagnosis and precise treatment of diseases. With the continuous advancement of deep learning technology and the continuous accumulation of medical imaging data, medical imaging diagnosis based on convolutional neural networks will play an increasingly important role in the future.

2. Related Work

In recent years, deep learning, particularly convolutional neural networks (CNNs), has played a crucial role in advancing medical image analysis. CNNs leverage multi-layer convolutional operations to automatically extract hierarchical features from images, which is particularly advantageous for processing complex medical data such as CT, MRI, and pathology images. Compared to traditional manual feature engineering, CNNs not only reduce the need for handcrafted features but also improve classification accuracy and generalization performance [8], [9]. Among CNN architectures, VGG16 has become one of the most widely applied models due to its deep yet simple structure, where small convolutional kernels (3×3) are stacked to enhance feature extraction capabilities. This design allows VGG16 to capture both local and global patterns, making it particularly effective for medical image classification tasks [10]. Beyond VGG16, ResNet-50 addresses the gradient vanishing problem that arises in deeper networks through its innovative residual learning mechanism, which preserves information flow across layers, making it suitable for complex medical image processing scenarios [11]. Inception-v3, with its inception modules designed to capture multi-scale spatial features simultaneously, provides flexibility in feature extraction, although its performance on highly specialized medical images often falls short of VGG16 and ResNet-50 [12]. MobileNet, designed for lightweight computation, offers advantages in mobile and embedded medical devices; however, its reduced capacity for feature representation can lead to lower classification accuracy in high-dimensional medical imaging tasks [13].

Apart from traditional CNN-based approaches, transformer models and attention mechanisms have emerged as complementary techniques for medical image analysis. Transformers, originally developed for natural language processing, have been successfully adapted for medical image classification, leveraging self-attention to capture long-range dependencies and complex spatial relationships within images [14]. Recent work incorporating multi-scale transformer architectures into medical image classification demonstrates the potential of these models to enhance performance by effectively integrating features across different resolution levels [15]. Furthermore, hybrid approaches that fuse CNNs with attention modules or transformers have shown promise in addressing data imbalance and fine-grained feature extraction challenges commonly encountered in medical imaging [16]. In addition to classification, deep learning models have been successfully applied to object detection tasks in medical images, such as lesion localization and organ segmentation. For instance, models like RT-DETR demonstrate the flexibility of combining region proposal mechanisms with transformer backbones, enhancing both detection accuracy and interpretability [17]. Adversarial robustness has also become a critical research topic, as adversarial attacks can severely undermine the reliability of medical image classifiers, exposing the need for robust training and adversarial defense techniques tailored to clinical applications [18].

To address the limited availability of labeled medical data, data augmentation and transfer learning are widely used strategies. Pre-trained CNNs such as VGG16 and ResNet-50 are commonly fine-tuned on medical image datasets, leveraging knowledge from large-scale natural image datasets to improve generalization on smaller medical image collections [19]. Beyond traditional transfer learning, techniques such as collaborative hypergraph networks have been explored for enhanced disease risk prediction and visit prediction in healthcare applications, demonstrating the potential of graph-enhanced deep learning for modeling complex patient interactions and medical histories [20], [21]. Knowledge distillation has also been employed to compress large models into smaller yet effective ones, addressing the need for efficient deep learning solutions in resource-constrained clinical environments [22]. Contrastive learning, particularly in cold-start scenarios where labeled data is scarce, has been applied to enhance feature representations in medical image classification [23].

Beyond medical image classification, deep learning has shown broad applicability across various medical and healthcare-related tasks, including survival prediction for cancer patients using deep neural networks, where multi-modal data integration plays a critical role [24]. Furthermore, convolutional networks have been extensively applied to cytopathology image classification, further demonstrating the flexibility and effectiveness of CNN-based approaches for domain-specific medical imaging tasks [25]. Outside the core medical imaging domain, concepts such as reinforcement learning for adaptive data mining and matrix logic approaches for efficient itemset discovery contribute to the broader methodological advancements that indirectly benefit healthcare analytics and medical data mining [26], [27]. Similarly, techniques such as federated learning are becoming increasingly important for scaling up medical vision-and-language models while preserving data privacy, particularly when integrating multi-institutional medical data [28]. Lastly, approaches such as spatiotemporal forecasting using hybrid deep learning techniques like LSTM combined with association rules show promise for broader healthcare system monitoring and prediction, which could complement image-based diagnostic models by providing contextual insights into patient status over time [29]. In summary, CNNs, transformers, and hybrid deep learning architectures form the backbone of modern medical image analysis, and continued advances in feature representation learning, model robustness, data-efficient learning, and privacy-preserving techniques will drive further progress in this critical domain.

3. Method

In this study, we used the convolutional neural network (CNN) model VGG16 for medical image classification tasks. The VGG16 model automatically extracts image features through multi-layer convolution and pooling operations, and then classifies through fully connected layers. In order to describe the workflow and reasoning process of VGG16 in detail, the structure of the model and its application in medical image classification will be introduced below. Its model architecture is shown in Figure 1.

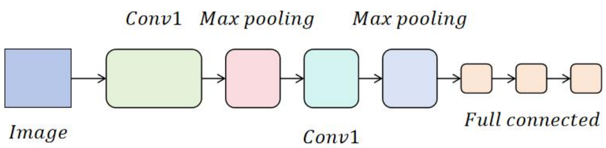

Figure 1. Model network architecture

The basic structure of VGG16 includes 16 layers, 13 of which are convolutional layers and 3 are fully connected layers. Each convolution operation extracts features from the input image, using a small 3 × 3 convolution kernel, and the convolution stride is usually 1, while the pooling operation uses a 2 × 2 maximum pooling layer for downsampling to reduce the dimension of the feature map. Assume that the input image is A, where H represents the height of the image, W represents the width of the image, and C represents the number of channels. The convolution operation can be expressed as:

$$Y = Conv(X) = f(X * W + b)$$

Among them, W is the convolution kernel, b is the bias, * represents the convolution operation, f is the activation function, and Y is the output feature map after the convolution operation.

In VGG16, the convolutional layer is usually followed by a maximum pooling layer, which performs downsampling through a 2X2 pooling window. The pooling operation can be expressed as:

$$Y_p = Pool(Y)$$

The pooling operation reduces the spatial resolution of the feature map by selecting the maximum value in the pooling window, thereby effectively reducing the computational complexity. The pooled feature map will be input to the next convolutional layer for further processing.

The last few layers of VGG16 are fully connected layers, which are used to classify based on the features extracted previously. Assume that the output of the previous layer is a feature vector $z \in R^N$, where N is the dimension of the feature vector. The output of the fully connected layer is calculated by the following linear transformation:

$$y = W_f z + b_f$$

Among them, W_f is the weight matrix of the fully connected layer, b_f is the bias, and y is the output vector of the fully connected layer. Finally, we convert the output vector into category probability through the softmax function:

$$P(y = k \mid X) = \frac{e^{yk}}{\sum_{i=1}^{K} e^{y_i}}$$

Where y_k is the k-th element in the output vector of the fully connected layer, K is the number of categories, and $P(y = k \mid X)$ is the probability that the input image belongs to the k-th category.

During model training, we use the cross entropy loss function to optimize the parameters of the network. Assuming that the training set contains N samples, y_i is the true label of the i-th sample, and B is the probability distribution predicted by the network, the cross entropy loss function can be expressed as:

$$L = -\frac{1}{N} \sum_{i=1}^{N} \sum_{k=1}^{K} y_{ik} \log(y'_{ik})$$

Among them, y_{ik} is the value of the true label of sample i in the k-th class, and y'_{ik} is the predicted probability of sample i in the k-th class. By minimizing the cross entropy loss function, the network can continuously adjust the weights and biases to improve the classification accuracy.

In order to avoid overfitting and improve the generalization ability of the model, this study adopted the strategies of data augmentation and transfer learning. Data augmentation technology expands the diversity of the data set by rotating, translating, scaling and other operations on the training data. Transfer learning uses the VGG16 model pre-trained on a large-scale image dataset and transfers it to the medical image classification task, thereby accelerating the training process and improving the performance of the model.

In general, the VGG16 model can effectively extract high-level features from medical images and accurately classify them through the structure of multi-layer convolution, pooling and fully connected layers. By optimizing the model's loss function and combining data enhancement and transfer learning, the model can achieve better performance in medical image classification tasks.

4. Experiment

4.1. Datasets

This study uses a public medical imaging dataset, Chest X-ray14, which contains chest X-ray images from 14 different types of diseases and is widely used for the detection and classification of lung diseases. The dataset contains about 100,000 chest X-ray images with an image size of 1024 × 1024 pixels, covering a variety of lung diseases, including pneumonia, tuberculosis, lung cancer, acute bronchitis, etc. Each image is equipped with a corresponding label to indicate whether there are relevant lesions and disease types in the image. This dataset is provided by the National Cancer Institute (NCI) of the United States. It is a relatively comprehensive and challenging medical imaging dataset suitable for the training and verification of deep learning models.

Each image in the dataset has been annotated by professional doctors to ensure the accuracy and reliability of the labels. The label of each case image contains multiple binary classification labels, indicating whether a specific disease exists. Therefore, this dataset is not only suitable for the classification problem of multiple diseases, but also can be used for multi-label classification tasks. The image file format is JPEG and has been standardized to ensure the stability and efficiency of model training. The scale of the dataset and the detailed annotation make it one of the commonly used and standard evaluation datasets in the field of medical image analysis.

To ensure the diversity and generalization ability of model training, the images in the dataset have been subjected to various data enhancement processes, such as random cropping, rotation, flipping, etc., which enables the model to adapt to different image transformations during training and improves its robustness. Through the training of this dataset, the VGG16 model can effectively learn the key features in chest X-ray images and show a high accuracy in the diagnosis of lung diseases. In addition, due to the certain scale and representativeness of the dataset, its training results can provide valuable reference for actual clinical diagnosis.

4.2. Experimental Results

In order to verify the effectiveness of the proposed VGG16 model in medical image classification, we conducted comparative experiments with four classic image classification models. First, we selected the traditional convolutional neural network (CNN) as the baseline model to evaluate its basic performance in medical image classification. Secondly, we adopt the ResNet-50 model, which alleviates the vanishing gradient problem in deep networks by introducing residual connections and has stronger feature learning capabilities. Then, we used the Inception-v3 model, which can capture rich features in images through the design of multi-scale convolution kernels and is suitable for complex medical image analysis tasks. Finally, we also chose a lightweight model - MobileNet. MobileNet significantly reduces the number of parameters of the model through depth-separable convolution, making it suitable for application scenarios that require low computing resources. All models were trained under the same data set and training conditions, and their performance was comprehensively compared using three evaluation indicators: accuracy, recall, and F1-Score. The experimental results are shown in Table 1.

Table 1. Experimental Results

Model	ACC	Recall	F1-Score
CNN	82.4	80.3	81.3
Inception-v3	86.5	84.2	85.3
MobileNet	84.7	82.5	83.6
ResNet-50	88.0	86.1	87.0
VGG16	90.2	88.5	89.3

It can be seen from the experimental results that VGG16 performed the most outstandingly in all evaluation indicators, with a precision (ACC) of 90.2%, a recall rate (Recall) of 88.5%, and an F1 score of 89.3%. VGG16 can effectively extract useful features from medical images through its deep network structure and hierarchical feature extraction capabilities, especially its high performance in classification accuracy and recall, indicating that it can handle complex medical image tasks. It has strong robustness and accuracy. Its advantage is mainly reflected in the fact that deep convolutional networks can identify key features of diseases in small differences in images by learning and refining image features layer by layer, significantly improving the classification performance of the model. Therefore, the performance of VGG16 in medical imaging diagnosis tasks provides an excellent reference standard for other deep learning models.

Compared with VGG16, ResNet-50 also showed higher performance in this experiment, with a precision of 88.0%, a recall rate of 86.1%, and an F1 score of 87.0%. ResNet effectively solves the vanishing gradient problem in deep network training by introducing residual connections, allowing deeper network structures to better capture complex features in images. This gives ResNet certain advantages in medical image classification, especially when processing more complex images, it can ensure that the information flow is effectively propagated in the network, thereby improving classification performance. However, although ResNet's performance is better, it is still slightly inferior to VGG16. This may be due to the higher complexity of its network architecture, which may take longer to tune during the training process, and has higher requirements for data diversity.

Inception-v3 and MobileNet performed slightly worse than VGG16 and ResNet-50 in this experiment. The precision of Inception-v3 is 86.5%, the recall rate is 84.2%, and the F1 score is 85.3%; the precision of MobileNet is 84.7%, the recall rate is 82.5%, and the F1 score is 83.6%. Through its multi-scale convolution operation, Inception-v3 can capture a variety of features of the image, especially in the details of the image, and can improve the robustness to complex backgrounds. However, despite its advantages of multi-scale convolution, Inception-v3 still performs worse than VGG16 and ResNet-50 in precision and recall, possibly due to the fact that this model requires more computing resources and training time to achieve ideal classification. Effect. As a lightweight model, MobileNet is mainly optimized for devices with limited computing resources. Although it performs well on mobile devices, its lower precision and recall show its limitations in processing complex medical images. Capabilities are limited. Although MobileNet has optimized the number of parameters and calculations to improve the running speed, its performance is relatively average in actual medical image classification tasks.

The CNN model performed the worst among all compared models, with a precision of 82.4%, a recall rate of 80.3%, and an F1 score of 81.3%. This result shows that although CNN, as a basic model in the field of deep learning, can effectively extract image features, it has certain limitations when facing complex medical images. The structure of CNN is relatively simple and lacks deeper feature learning capabilities and more complex network structures to cope with complex patterns in medical image classification. Its lower precision and recall indicate that CNN cannot extract useful features as accurately as deeper networks (such as VGG16 and ResNet-50) in the recognition of small differences, complex backgrounds and irregular shapes in medical images, resulting in Its performance on this task is far inferior to other more complex models.

Overall, both VGG16 and ResNet-50 perform well in medical image classification tasks, with VGG16 showing particular strength. Its deeper convolutional layers and robust feature extraction capabilities enable it to capture more detailed information, leading to higher classification accuracy and recall rates. Although Inception-v3 and MobileNet each have their own advantages — Inception-v3 enhances feature learning diversity through multi-scale convolution, while MobileNet improves computational speed with its lightweight design — neither can match the performance level of VGG16 and ResNet-50 in complex medical imaging tasks. In addition, although the CNN model is basic, its relatively simple structure and low performance limit its application in actual medical imaging diagnosis. Therefore, based on the experimental results, it can be speculated that VGG16 and ResNet-50 are ideal choices for medical image classification tasks, especially in tasks that require high precision and high recall.

5. Conclusion

This study draws some key conclusions through experimental comparison of five models, VGG16, ResNet-50, Inception-v3, MobileNet and CNN, in medical image classification tasks. The experimental results show that VGG16 outperforms other models in precision, recall and F1 score, proving its superior performance in complex medical image analysis. ResNet-50 follows closely behind, effectively alleviating the gradient vanishing problem in deep networks with its residual structure, and also showing good classification performance. Although Inception-v3 and MobileNet are innovative in model design, their performance in processing complex medical images is relatively weak, especially in recall and F1 score, which fail to reach the level of VGG16 and ResNet-50. As a baseline model, CNN has low classification accuracy and recall when facing medical images, indicating that a deeper network structure is crucial to improving classification performance.

Future research can further explore how to optimize existing convolutional neural networks, especially in medical image analysis, how to improve the accuracy and robustness of the model by combining multiple network architectures (such as the combination of VGG16 and ResNet). In addition, with the continuous increase of medical image data, how to use data enhancement, transfer learning, and self-supervised learning techniques to further improve the generalization ability of the model in small sample data sets is also a direction worthy of in-depth discussion. With the improvement of computing power and innovation in model design, deep learning will play an increasingly important role in the field of medical image analysis, especially in early disease detection and precise diagnosis and treatment, and has broad application prospects.

References

[1] Mascarenhas S, Agarwal M. A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for Image Classification[C]//2021 International conference on disruptive technologies for multi-disciplinary research and applications (CENTCON). IEEE, 2021, 1: 96-99.

[2] Pugliesi R A. Deep Learning Models for Classification of Pediatric Chest X-ray Images using VGG-16 and ResNet-50[J]. Sage Science Review of Applied Machine Learning, 2019, 2(2): 37-47.

[3] Shivadekar S, Kataria D B, Hundekar S, et al. Deep learning based image classification of lungs radiography for detecting COVID-19 using a deep CNN and ResNet 50[J]. International Journal of Intelligent Systems and Applications in Engineering, 2023, 11: 241-250.

[4] Zakaria N, Mohamed F, Abdelghani R, et al. VGG16, ResNet-50, and GoogLeNet deep learning architecture for breathing sound classification: a comparative study[C]//2021 International Conference on Artificial Intelligence for Cyber Security Systems and Privacy (AI-CSP). IEEE, 2021: 1-6.

[5] Wen L, Li X, Gao L. A transfer convolutional neural network for fault diagnosis based on ResNet-50[J]. Neural Computing and Applications, 2020, 32(10): 6111-6124.

[6] Aatila M, Lachgar M, Hrimech H, et al. Diabetic retinopathy classification using ResNet50 and VGG-16 pretrained networks[J]. International Journal of Computer Engineering and Data Science (IJCEDS), 2021, 1(1): 1-7.

[7] Khan H A, Jue W, Mushtaq M, et al. Brain tumor classification in MRI image using convolutional neural network[J]. Mathematical Biosciences and Engineering, 2021.

[8] Z. Zhu, Y. Zhang, J. Yuan, W. Yang, L. Wu, and Z. Chen, "NLP-driven privacy solutions for medical records using Transformer architecture," unpublished.

[9] Z. Gao, T. Mei, Z. Zheng, X. Cheng, Q. Wang, and W. Yang, "Multi-channel hypergraph-enhanced sequential visit prediction," in Proc. Int. Conf. Electron. Devices Comput. Sci. (ICEDCS), Sep. 2024, pp. 421-425.

[10] T. Mei, Z. Zheng, Z. Gao, Q. Wang, X. Cheng, and W. Yang, "Collaborative hypergraph networks for enhanced disease risk assessment," in Proc. Int. Conf. Electron. Devices Comput. Sci. (ICEDCS), Sep. 2024, pp. 416-420.

[11] L. Wu, J. Gao, X. Liao, H. Zheng, J. Hu, and R. Bao, "Adaptive attention and feature embedding for enhanced entity extraction using an improved BERT model," unpublished.

[12] X. Yan, W. Wang, M. Xiao, Y. Li, and M. Gao, "Survival prediction across diverse cancer types using neural networks," in Proc. 7th Int. Conf. Mach. Vis. Appl., 2024, pp. 134-138.

[13] M. Xiao, Y. Li, X. Yan, M. Gao, and W. Wang, "Convolutional neural network classification of cancer cytopathology images: taking breast cancer as an example," in Proc. 7th Int. Conf. Mach. Vis. Appl., 2024, pp. 145-149.

[14] Y. Yang, "Adversarial attack against image classification based on generative adversarial networks," arXiv:2412.16662, 2024.

[15] S. Wang, C. Wang, J. Gao, Z. Qi, H. Zheng, and X. Liao, "Feature alignment-based knowledge distillation for efficient compression of large language models," arXiv:2412.19449, 2024.

[16] F. Shao, T. Zhang, S. Gao, Q. Sun, and L. Yang, "Computer vision-driven gesture recognition: Toward natural and intuitive human-computer interaction," arXiv:2412.18321, 2024.

[17] X. Huang, Z. Zhang, X. Li, and Y. Li, "Reinforcement learning-based Q-learning approach for optimizing data mining in dynamic environments," unpublished.

[18] Y. Li, W. Zhao, B. Dang, X. Yan, M. Gao, W. Wang, and M. Xiao, "Research on adverse drug reaction prediction model combining knowledge graph embedding and deep learning," in Proc. 4th Int. Conf. Mach. Learn. Intell. Syst. Eng. (MLISE), Jun. 2024, pp. 322-329.

[19] X. Li, T. Ruan, Y. Li, Q. Lu, and X. Sun, "A matrix logic approach to efficient frequent itemset discovery in large data sets," arXiv:2412.19420, 2024.

[20] B. Chen, F. Qin, Y. Shao, J. Cao, Y. Peng, and R. Ge, "Fine-grained imbalanced leukocyte classification with global-local attention transformer," J. King Saud Univ. - Comput. Inf. Sci., vol. 35, no. 8, Article ID 101661, 2023.

[21] X. Sun, Y. Yao, X. Wang, P. Li, and X. Li, "AI-driven health monitoring of distributed computing architecture: Insights from XGBoost and SHAP," arXiv:2501.14745, 2024.

[22] J. Hu, T. An, Z. Yu, J. Du, and Y. Luo, "Contrastive learning for cold start recommendation with adaptive feature fusion," arXiv:2502.03664, 2025.

[23] J. Song and Z. Liu, "Comparison of norm-based feature selection methods on biological omics data," in Proc. 5th Int. Conf. Adv. Image Process., Nov. 2021, pp. 109-112.

[24] J. Hu, Y. Xiang, Y. Lin, J. Du, H. Zhang, and H. Liu, "Multi-scale Transformer architecture for accurate medical image classification," arXiv:2502.06243, 2025.

[25] X. Liao, C. Wang, S. Zhou, J. Hu, H. Zheng, and J. Gao, "Dynamic adaptation of LoRA fine-tuning for efficient and task-specific optimization of large language models," arXiv:2501.14859, 2025.

[26] W. He, Y. Zhang, T. Xu, T. An, Y. Liang, and B. Zhang, "Object detection for medical image analysis: Insights from the RT-DETR model," arXiv:2501.16469, 2025.

[27] J. Wang, "Markov network classification for imbalanced data with adaptive weighting," J. Comput. Sci. Softw. Appl., vol. 5, no. 1, pp. 43-52, 2025.

[28] S. Lu, Z. Liu, T. Liu, and W. Zhou, "Scaling-up medical vision-and-language representation learning with federated learning," Eng. Appl. Artif. Intell., vol. 126, 107037, 2023.

[29] Y. Deng, "A hybrid network congestion prediction method integrating association rules and LSTM for enhanced spatiotemporal forecasting," Trans. Comput. Sci. Methods, vol. 5, no. 2, 2025.