Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 9
AIDML 2025

Development Trends and Future Prospects of Artificial Intelligence in Medical Image Processing

Cheng Luo *

Maynooth International Engineering College, Fuzhou University, Fuzhou, China

* Corresponding Author Email: 832203323@fzu.edu.cn

Abstract. In recent years, the momentum of the development of artificial intelligence has become increasingly strong, and the integration of artificial intelligence and various disciplines has also been applied to all aspects of human life. At the same time, in the medical industry related to human health, research on artificial intelligence + medicine has been carried out for many years. In the medical field, the processing of medical images already accounts for 80%-90% of the medical data sources, playing an extremely important role in the doctor's diagnosis and treatment plan formulation. However, a large amount of medical imaging data such as X-ray imaging, CT, MRI, etc. is handed over to doctors to manually identify disease lesions, which inevitably leads to some errors. In addition, the annual growth rate of the above medical images is roughly 30%, but the annual growth rate of radiologists is only 4%, which makes image analysis doctors face more severe analysis tasks. To address such needs, multiple models that apply deep learning technology to medical image analysis have emerged. This article will focus on analyzing the application of three deep learning models in the field of medical image analysis: convolutional neural networks (CNNs), generative adversarial networks (GANs), U-Net and its variants, and introduce their latest development status, analyze and compare their advantages and disadvantages, analyze the technical challenges that the model will encounter in practical applications, and make reasonable future prospects.

Keywords: AI; Medical Imaging; CNN; GAN; U-Net.

1. Introduction

With the rapid development of human society, people are paying more attention to their own physical health, which also promotes the rapid development of medicine. In modern medicine, various imaging technologies have become important references to assist doctors in diagnosis, such as computed tomography (CT) and magnetic resonance imaging (MRI). However, the images obtained by medical imaging technology still need to go through the step of medical imaging processing before they can provide doctors with clear images of the disease. However, the growth in the number of radiologists cannot keep up with the increasing amount of medical impact data. This medical situation has led to the explosive application of artificial intelligence technology in the field of medical image processing.

Medical image processing covers multiple task types, including image enhancement, organ segmentation, and lesion detection. If traditional processing methods are used, it will face three major challenges: difficulty in identifying tiny lesions, differences in diagnostic results (the diagnostic consistency of different radiologists for the BI-RADS classification of breast mammography is only 68% [1]), and low efficiency in 3D image processing. AI technology helps innovate medical image processing technology through technical features such as the local perception capabilities of convolutional neural networks (CNN) and the image enhancement capabilities of generative adversarial networks (GAN).

As AI technology demonstrates its powerful processing capabilities in medical image processing, research on AI technology in this field has become more in-depth and comprehensive, and many important results have emerged. There have been great breakthroughs in the field of basic models of deep learning. This article summarizes the recent research status of various deep learning models:

Qi Han et al. proposed a new CNN model called DM-CNN (Dynamic Multi-scale Convolutional Neural Network), which focuses on medical image classification tasks and achieved classification accuracy rates of 97. 30%, 99. 89%, 98. 31%, and 98. 97% on four medical imaging datasets, and integrated the uncertainty quantification function [2]. Hongwei Ding and others invented LEGAN. Experiments on medical imaging datasets (BloodMNIST, OrgancMNIST, PathMNIST) showed that the model has significant advantages over other models: High generation quality: On BloodMNIST, LEGAN's FID (Fréchet Inception Distance) is 10. 80 and ppl (Perceptual Path Length) is 0. 27, which is more than 40% lower than the baseline model. Good classification performance: 92. 1% accuracy on PathMNIST, strong intra-class coverage: only 5% of rare cell morphologies are found in PathMNIST. By integrating sparsity and entropy constraint mechanisms, it provides a new paradigm with high generalization for small sample enhancement in medicine [3]. Since Ronneberger et al. proposed the U-Net model in 2015 [4], the emergence of U-Net has effectively solved the model challenges of variable object scales and blurred boundaries in medical images. However, the traditional U-Net model has obvious defects in modeling the sequence relationship between layers: the deep network loses detail information due to step-by-step downsampling, and when the jump connection directly splices the shallow and deep features, the fusion effect is limited due to the semantic gap. Although many variants have been developed to address this problem, these variants generally have the core bottleneck of insufficient utilization of historical information between levels and unbalanced expression of multi-band features. Facing this challenge, I^2U-Net has achieved a breakthrough through dual-path architecture and multi-functional interaction mechanism. Its image feature path (improved ResNet-34 encoder) retains multi-scale details, the hidden state path (zero-initialized input) simulates RNN time series modeling, and the MFII module realizes cross-path and cross-layer information interaction;The HIFA module combines local convolution and non-local attention to extract high/low frequency features in a balanced manner. Experiments show thatI^2U-Net significantly outperforms traditional U-Net variants in tasks such as skin lesions (ISIC2018 IoU 83. 66%) and polyps (average Dice 82. 13%), providing a more robust solution for medical image segmentation [5]. There are also significant research results for deep learning models such as recurrent neural networks (RNN), long short-term memory networks (LSTM), and transformers. For example, the enhanced RNN-LSTM model proposed by H. N. Veena et al. has an accuracy of 97. 4% on the DRISHTI-GS dataset, which is significantly better than the traditional U-Net model (95%) and SVM method (89%). Its sensitivity and specificity are also improved to 97% and 97. 9%, respectively [6]. The SSTrans-Net proposed by Liyao Fu et al. achieved an average Dice score of 82. 89% (3. 76% higher than SwinUNet) on the Synapse medical imaging dataset with only 29. 64M parameters, outperforming other mainstream models [7].

This article will systematically review the latest development trends and prospects of artificial intelligence in the field of medical image processing, focusing on the breakthrough innovations of many deep learning models and their clinical experimental applications. This article will also explore in depth the experimental performance of new deep learning models developed based on traditional deep learning models in various medical image processing tasks, and compare multiple deep learning models to show the advantages and disadvantages of different models as well as future challenges.

2. Typical Technology Comparison and Analysis

2.1. Convolutional Neural Networks (CNNs)

2.1.1. Technical principle

Convolutional neural network (CNN) is a deep learning model based on local perception and weight sharing. Its core architecture includes convolutional layer, pooling layer and fully connected layer. The convolutional layer extracts local features through sliding windows (convolution kernels), the pooling layer achieves feature dimensionality reduction and translation invariance, and the fully connected layer completes the classification task. In medical image processing, CNN can capture the morphological features (such as edges, textures, and density differences) of lesions from complex images such as X-rays, CT, and MRI through its multi-level feature abstraction capabilities.

2.1.2. Research status

In recent years, the improvement of CNN for medical image classification tasks has mainly focused on the fields of multi-scale feature fusion and uncertainty quantification. For example, the dynamic multi-scale convolutional neural network (DM-CNN) with uncertainty quantification proposed by Qi Han et al. achieved breakthroughs through the following innovations. Dynamic multi-scale feature fusion (DMFF) adaptively selects convolution kernels of different sizes such as 3×3, 5×5, and 7×7 according to the feature level to extract multi-granular features (such as the local texture and overall morphology of skin lesions). Its purpose is to find a balance between network performance and computing load, and to integrate high-level semantic information and low-level semantic information to enhance feature representation capabilities. Hierarchical Uncertainty Quantized Attention (HDUQ-Attention) introduces the Monte Carlo (MC) Dropout structure into the attention mechanism. By dynamically adjusting the spatial attention and channel attention weights of features at different levels, the model's ability to focus on key areas of medical images, such as lung nodule localization, is improved. At the same time, the probability distribution histogram can intuitively display the prediction confidence, which can significantly assist doctors in their judgment; The innovation of Multi-scale Fusion Pooling (MF Pooling) lies in designing a pooling strategy for multi-scale models. Through cross-layer feature interaction and adaptive pooling levels, cross-layer feature interaction information is retained during dimensionality reduction, thereby enhancing the robustness of the model to multi-scale manifestations of lesions and reducing the computational complexity of the model. The innovation of the multi-objective loss function (MO Loss) module lies in the design of a differentiated loss function combination based on the difference in parameter quantities between the backbone network and other modules of the model. SmoothL1 loss is used to calculate the loss of the backbone network, which is robust to outliers. Cross entropy loss is used to strengthen the classification boundaries when calculating the losses of other modules. Finally, the loss of the weighted fused model is calculated through the module parameter ratio to alleviate the optimization bias caused by the imbalance of module parameters and accelerate model convergence. In experimental performance, the model significantly outperforms traditional architectures such as ResNet-50 and DenseNet-121 in tasks such as ophthalmology (glaucoma classification) and respiratory medicine (pneumonia detection). The experimental performance of the model in the experiment is shown in Figures 1, 2, 3 and 4. Its uncertainty quantification function provides a credible reference for clinical decision-making [2].

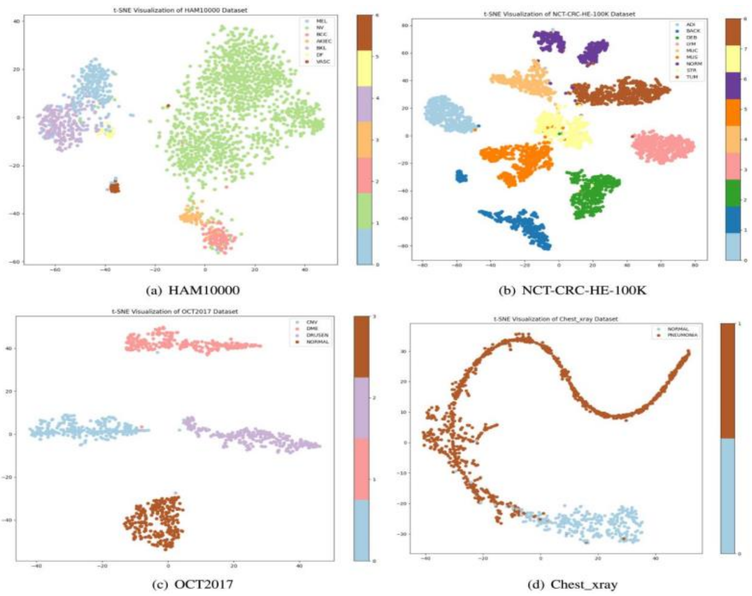
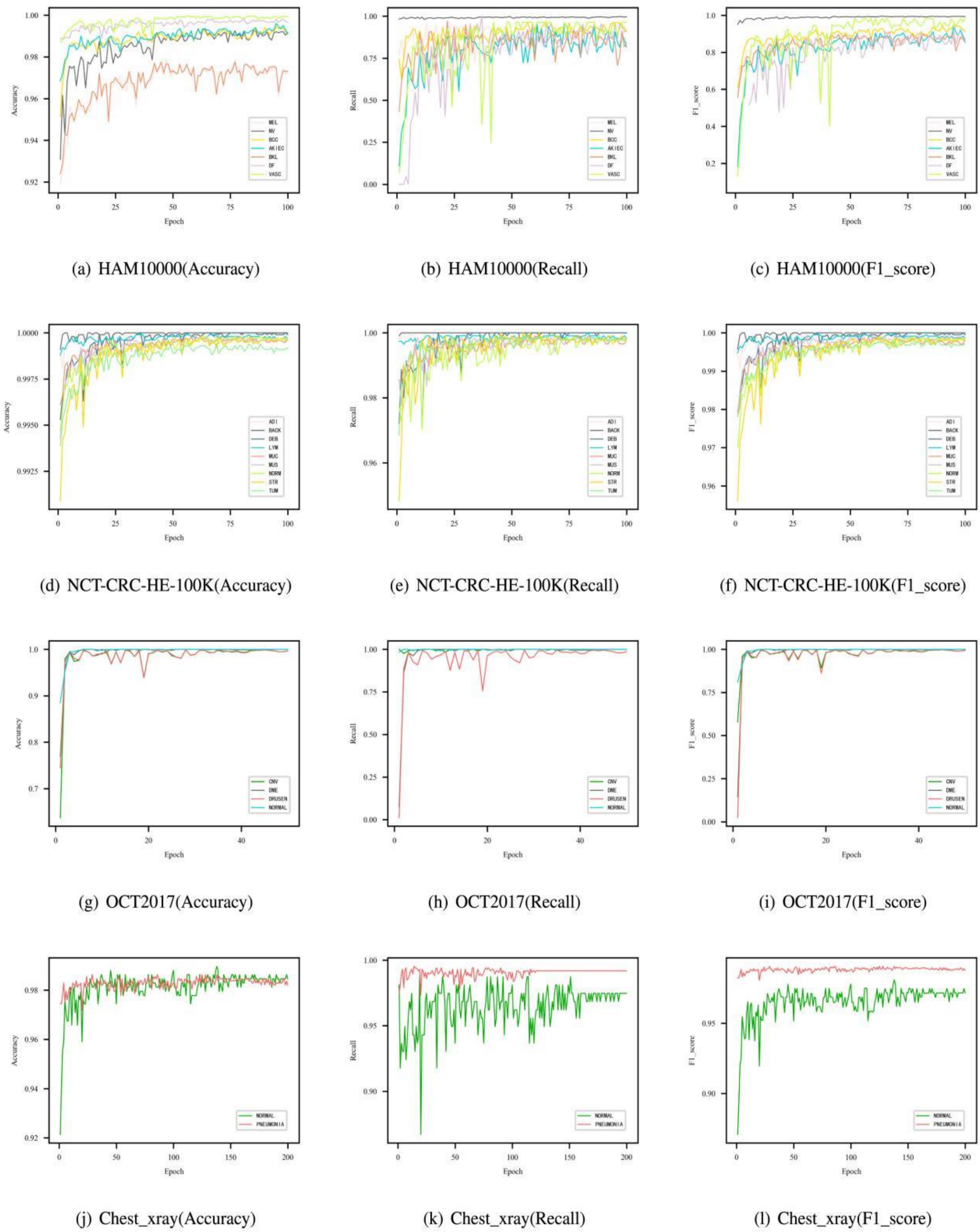


Figure 1. t-SNE visualization of DM-CNN (ResNet101) on four datasets (HAM10000, NCT-CRC-HE-100K, OCT2017, Chest_xray) [2].



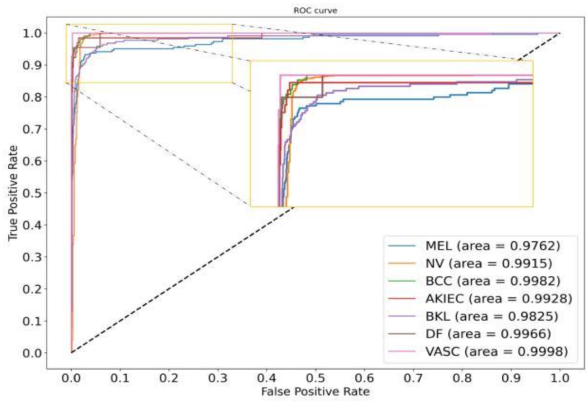
(a) HAM10000(Accuracy) (b) HAM10000(Recall) (c) HAM10000(F1_score)

(d) NCT-CRC-HE-100K(Accuracy) (e) NCT-CRC-HE-100K(Recall) (f) NCT-CRC-HE-100K(F1_score)

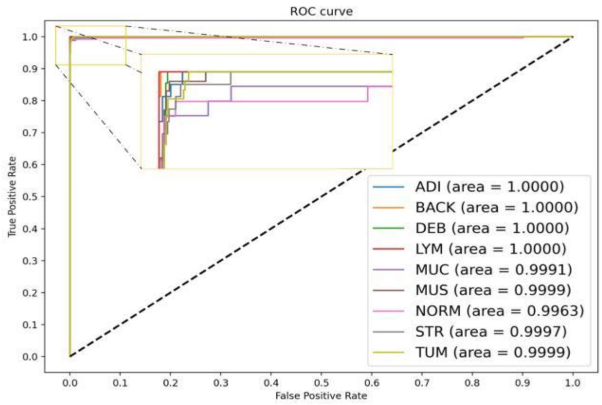
(g) OCT2017(Accuracy) (h) OCT2017(Recall) (i) OCT2017(F1_score)

(j) Chest_xray(Accuracy) (k) Chest_xray(Recall) (l) Chest_xray(F1_score)

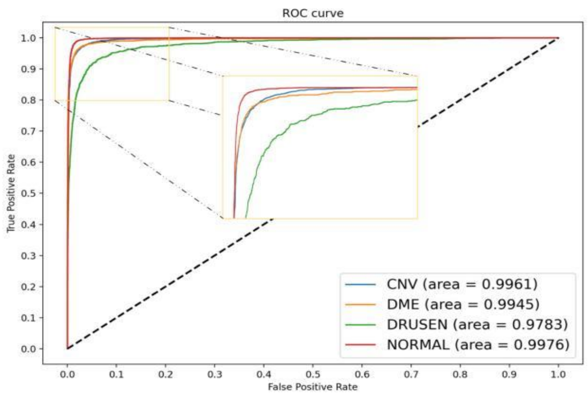
Figure 2. The detailed performance indicators of the proposed DM-CNN model are tested on four datasets (HAM10000, NCT-CRC-HE-100K, OCT2017, Chest_xray) [2].



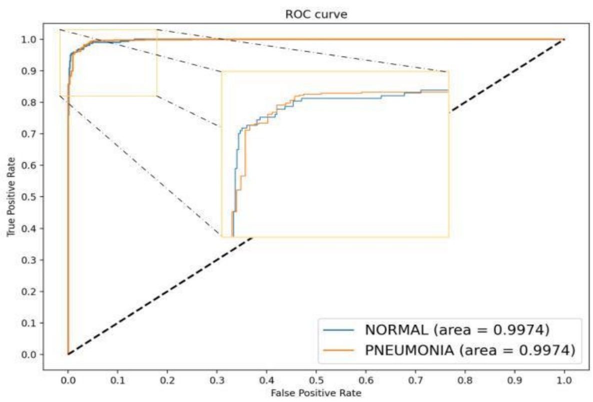
(a) HAM10000



(b) NCT-CRC-HE-100K

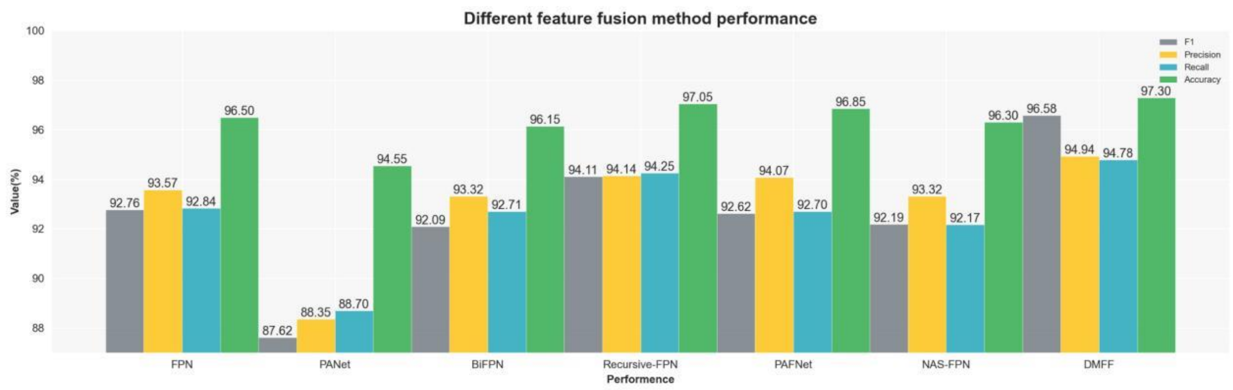


(c) OCT2017

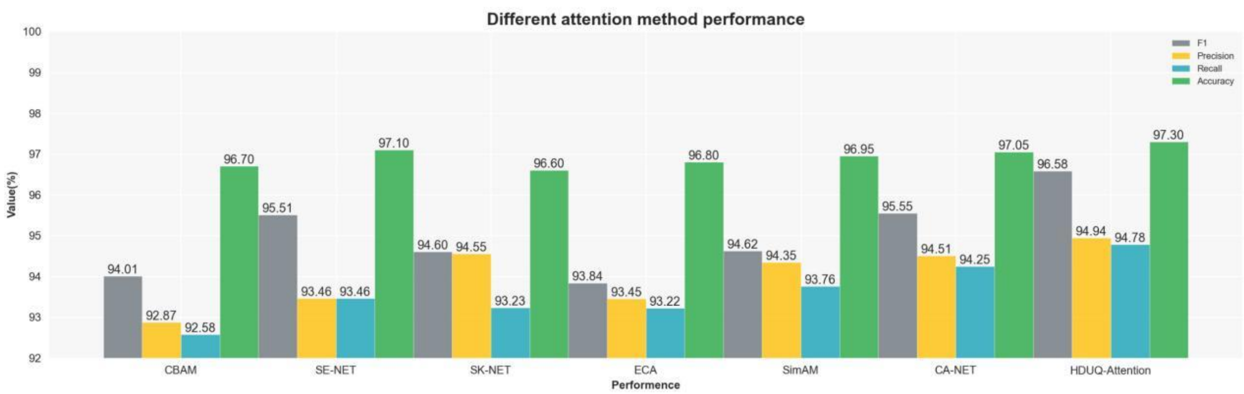


(d) Chest_xray

Figure 3. ROC curves of DM-CNN (ResNet101) on four datasets (HAM10000, NCT-CRC-HE-100K, OCT2017, Chest_xray) [2].



(a)



(b)

Figure 4. (a) Performance histogram of DMFF compared with other fusion methods; (b) Performance histogram of HDUQ-Attention compared with other attention methods [2]

2.1.3. Advantages and disadvantages analysis

Convolutional neural networks (CNNs) have shown significant advantages in medical text classification tasks. It has a relatively powerful automatic special diagnosis extraction capability, and can automatically learn local features in images through convolution kernels without relying on manually designed features. It is suitable for automated analysis of complex lesions, such as brain tumor image processing and breast cancer classification. CNN also has the ability to process high-dimensional data. It can effectively process high-resolution medical images, such as full-slice case images, 3D MRI, etc., through pooling layers and downsampling operations, reducing computational complexity. For example, Seitaro Baba et al. proposed using 3DCNN to register head MRI images after awakening by incorporating global information. The results showed that the registration effect was significantly improved [8], Convolutional neural networks also have the advantages of high efficiency in transfer learning applications to solve the problem of insufficient data; they also have end-to-end learning capabilities, which can reduce the occurrence of information loss in traditional processes; and their diagnostic accuracy exceeds that of senior physicians.

Although CNN models perform well, they still have some limitations. First, CNN is highly dependent on the amount and quality of data, requiring professional physicians to participate in the annotation of medical images, and small sample data can easily lead to overfitting; Secondly, due to the "black box" characteristics of the CNN model, the model has poor interpretability, and doctors find it difficult to trust the decision logic of the model, and cannot trace the source of misdiagnosis; Finally, CNN also faces challenges such as high demand for computing resources and limited generalization ability.

2.2. Generative Adversarial Networks (GANs)

2.2.1. Technical Principle

Generative Adversarial Network (GAN) is a deep learning model that generates data through adversarial training. The core technical principle of this model is to allow the generator and the discriminator to evolve together while playing a game, and finally generate data that is more realistic. GAN consists of two neural networks, the generator and the discriminator. Random noise is input into the generator and forged data, such as images, are output. The discriminator judges the output of the generator. After inputting real data or forged data generated by the generator, the discriminator outputs a probability value to indicate the probability that the input data is real. The two reach a balance in adversarial training, making the distribution of generated data almost consistent with that of real data.

2.2.2. **Research** Status

The powerful data enhancement and sample expansion capabilities demonstrated by GAN can be effectively used in the processing of medical images. Especially when medical data is scarce or the cost of obtaining medical data is high, the ability of GAN can be greatly utilized to generate synthetic images that are highly similar to real data for doctors to diagnose. In other aspects, GAN can also improve image resolution, image denoising, lesion segmentation and other functions to assist doctors in diagnosis. However, GAN is also prone to Intraclass Mode Collapse when generating medical images. The generated samples only cover the dense areas within the category and ignore the sparse areas, exacerbating data bias. LEGAN, proposed by Hongwei Ding et al., made innovations to this challenge. The model locates sparse samples based on the LOF algorithm, and then enhances the sparse samples through affine transformation (AT), forcing the generator to learn sparse samples, LEGAN uses the information entropy formula to evaluate whether the generated samples are evenly distributed in dense and sparse areas. At the same time, the entropy value is added as a regular term to the generator's loss function, forcing the generator to optimize in the direction of maximizing diversity, At the same time, LEGAN introduces L1 reconstruction loss (Lr) to constrain the similarity between the generated samples and the original samples to ensure that the images generated by LEGAN are realistic and appropriate. LEGAN was verified on three medical image datasets

(BloodMNIST, OrgancMNIST, PathMNIST) and showed good performance in production quality, classification performance, and visual verification, as shown in Figure 5.

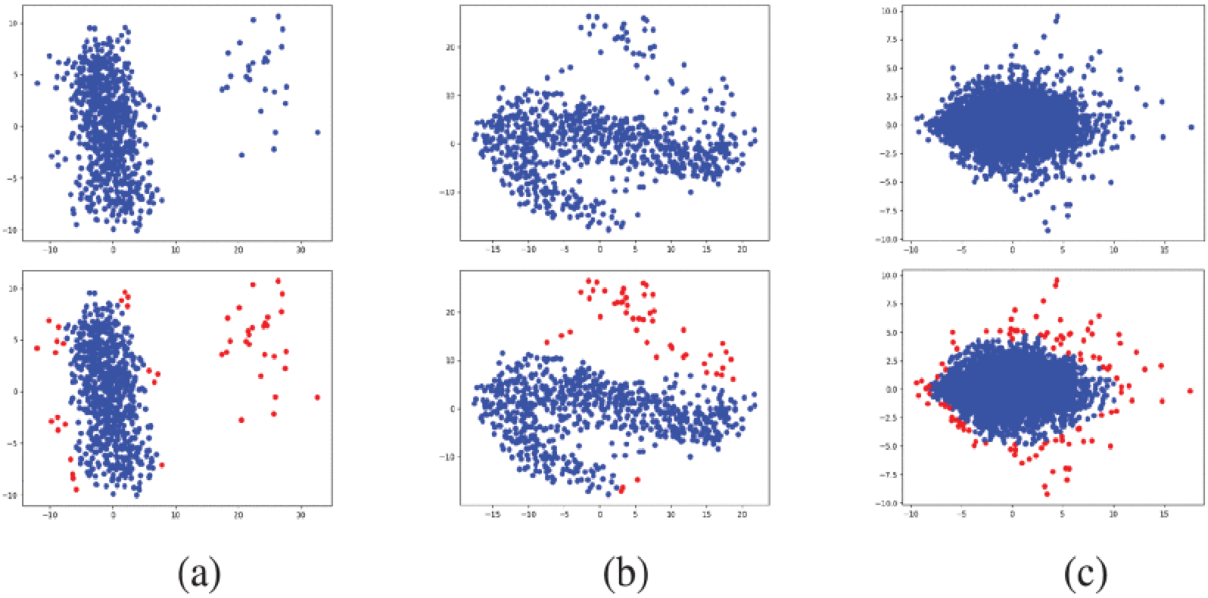


Figure 5. Sparse point detection results based on LOF.
(a) In-class detection results of BloodMNIST data,
(b) In-class detection results of OrgancMNIST data, and
(c) In-class detection results of PathMNIST data [9]

Using GAN in the field of medical image processing can indeed bring out its ability to generate synthetic data, but there is a risk of training data leakage, which may lead to the privacy of some patients being violated. In response to this, Hansle Gwon et al. proposed a new GAN model called LDP-GAN, which combines local differential privacy (LDP) with GAN. By using LDP noise on the original medical data, LDP directly perturbs the data itself instead of the gradient like the traditional global differential privacy (GDP) model to protect privacy, Feature-aware noise design and structure preservation were also performed. LDP-GAN also underwent multiple privacy attack tests, simulating various attack methods and quantifying the model's ability to resist attacks. The experimental results show that LDP-GAN is significantly better than other GDP-based models in terms of strong privacy protection [9].

2.2.3. Advantages and Disadvantages Analysis

The GAN model has many advantages in the field of medical image processing. For example, it can generate highly realistic medical images, medical texts, and electronic medical records through adversarial training. It can be well used in the training of disease prediction models and in data sharing under privacy protection, such as LDP-GAN [9]; GAN can also generate diverse samples by exploring different patterns of data distribution through random noise input to the generator, avoiding simple duplication of training data, GAN has the same unsupervised learning capability as CNN and can learn distribution from raw data without manual annotation. However, because GAN has the ability to generate realistic data, it is better suited for scenarios with smaller data sets than CNN.

The risks of the GAN model also exist. Due to the repeated adversarial training between the generator and the discriminator of the GAN, the generator may ignore certain disease characteristics, resulting in the diversity of real data being ignored by the GAN model. At the same time, adversarial training may also lead to unstable gradient updates, which need to be alleviated by relying on techniques, The quality of GAN-generated data is also difficult to evaluate. Quality assessment still relies on manual judgment or indirect indicators and is difficult to quantify directly.

2.3. U-Net and Its Variants

2.3.1. Technical Principle

U-Net is actually a convolutional neural network (CNN) designed specifically for medical image segmentation. U-Net is designed based on the characteristics of medical images, such as small data volume, fuzzy target boundaries, and the need for precise segmentation. It also optimizes the shortcomings of traditional CNN pooling operations that lead to loss of spatial information and difficulty in restoring detail positions, so that it can extract deep semantic features while retaining underlying detail information. U-Net is composed of an encoder and a decoder symmetrically. The encoder gradually extracts high-level semantic features and reduces the size of the feature map through convolution and pooling. The decoder gradually restores the size of the feature map through deconvolution and outputs the segmentation result. U-Net also uses skip connections so that the feature maps of each level of the encoder are cascaded with the corresponding decoder, which can integrate low-level details with high-level semantics.

2.3.2. Research Status

In recent years, due to the widespread application of U-Net in the field of medical image cutting and its extraordinary performance, research and innovation on U-Net have also been carried out vigorously, and many innovative models based on U-Net have emerged. Each innovative model has made a good improvement on the traditional U-Net. In the use of U-Net, doctors found that the traditional U-Net did not perform well in the segmentation of fuzzy boundaries and some irregular lesions in medical image segmentation. To address this problem, Duwei Dai et al. proposed a new dual-path U-Net---- I^2 U-Net, it introduces a dual-path structure, namely the image feature path (similar to the traditional U-Net) and the hidden state path (simulating the temporal memory of the recurrent neural network (RNN)), so that U-Net can effectively utilize historical information;I^2U-Net also introduces a multifunctional information interaction module (MFII) to support cross-path, cross-layer, and cross-path and layer interactions, and is designed to implement sequence modeling similar to recurrent neural networks (RNNs);At the same time, the overall information fusion and enhancement module (HIFA) is introduced to achieve the effect of multi-scale feature extraction, which can capture the global results and local edges of the organs at the same time and improve the segmentation accuracy of complex lesions. I^2U-Net was tested on 2D and 3D datasets, and its experimental performance was better than that of traditional U-Net and some existing U-Net variants (for example, the IoU in the skin lesion segmentation experiment reached 83. 66% and 85. 70%, significantly better than U-Net and Swin-Unet, and the average IoU in the polyp segmentation experiment reached 74. 91%, surpassing DenseASPP, H-Net and other methods) [5], Similarly, to address the problem that the traditional U-Net model is limited by local perception and has difficulty in modeling the global relationship between organs, Jieneng Chen et al. proposed TransUnet, which deeply integrates the Transformer Encoder and the traditional U-Net architecture. TransUnet flexibly integrates the Transformer's self-attention mechanism and cross-attention mechanism into the encoder and decoder of the traditional U-Net, allowing TransUnet to achieve a balance between global context modeling and local detail preservation [10][11], To address the accuracy problem of medical image segmentation, Guanqun Sun et al. proposed DA-TransUNet, which combines Transformer and dual attention mechanism (DA) to enhance the ability to extract spatial and channel features of medical images and improve feature transfer efficiency, significantly improve the model's sensitivity to subtle structures, and improve segmentation effects through the refined design of jump connections, There are also new U-Net models that are improved for different special scenarios, such as the CU-Net model for brain tumor MRI image segmentation proposed by Qimin Zhang et al. [12], which is used to improve the accuracy of tumor boundary demarcation and thus provide assistance to physicians in clinical treatment; For example, Pradip Senapati et al. proposed the Sharp Dense U-Net [13] to solve the problem of cell nucleus segmentation in histopathological images. It is used to assist pathologists in quickly and accurately locating cell nuclei, providing an automated tool for cancer diagnosis, especially in scenarios with limited resources.

2.3.3. Advantages and Disadvantages Analysis

Since the U-Net model was originally designed to serve the problem of medical image segmentation, its advantages in the field of medical image segmentation are quite obvious. U-Net can retain local features through data enhancement and skip connections to alleviate overfitting, which makes U-Net have efficient small sample learning ability and can achieve relatively good experimental performance with a small amount of labeled data, The symmetric encoder-decoder structure of U-Net also enables it to perform well in terms of spatial details and edge segmentation accuracy; and U-Net is flexible and scalable, and has spawned many variants, such as the previously mentioned U-Net [5], TransUnet [10], DA-TransUNet [11] and other innovative models based on U-Net.

U-Net is not a perfect model and it still has some flaws. For example, U-Net consumes a lot of computing resources. The number of parameters in a standard U-Net is about 7 million, and the jump connection of U-Net requires it to save multiple levels of special diagnosis images, which places high requirements on the video memory hardware. The U-Net model may also cause the loss of some tiny lesion information due to multiple downsampling, and the deconvolution or interpolation of U-Net may also cause it to generate blurred edges; since the jump connection of U-Net directly splices the features of the low-level and high-level layers, it may also cause fusion noise, resulting in the problem of semantic gap and feature mismatch.

3. Challenge

Although artificial intelligence technology has shown great development potential in medical image processing, its practical application still faces many technical difficulties, clinical and ethical challenges. First, in terms of the annotation and preprocessing of medical images, this work is highly dependent on the participation of professional physicians. Due to the differences in the physicians' supervisors, the annotation of images may also be different, which may lead to deviations in the model training results. At the same time, noise, low-contrast lesions in medical images, and alignment issues of multimodal data may affect the robustness of the model, Secondly, in terms of medical data privacy, since many medical institutions need to protect patients' private imaging data, these medical imaging data are difficult to share across institutions, which may lead to insufficient training data for the model, Then, in terms of human-computer collaboration, most of the existing artificial intelligence technologies now use "fully automatic" or "fully manual" modes, lacking dynamic human-computer interaction design, which may lead to misdiagnosis or missed diagnosis in the model diagnosis results, Finally, in terms of ethics, the definition of liability for medical disputes caused by AI misdiagnosis has not yet been clearly defined, there is a lack of laws to divide responsibilities, and the approval cycle for medical imaging AI models is an average of 3-5 years, which is much longer than the iteration speed of the model, which may result in the latest research results not being used in clinical treatment in a timely manner.

4. Conclusion

The rapid iterative development of artificial intelligence technology in the field of medical image processing has brought revolutionary changes to physicians' clinical diagnosis and treatment, and provided extremely useful auxiliary effects. This article introduces three types of deep learning models that are typical representatives in the field of medical image processing: convolutional neural network (CNN), generative adversarial network (GAN), and U-Net and its variants. It also systematically analyzes the basic technical principles, research status, technical advantages and limitations of these three core models, and reveals their breakthrough performance in medical image processing tasks such as medical image classification, image data enhancement and lesion segmentation. This shows that artificial intelligence technology has demonstrated important value in improving diagnostic efficiency and reducing physicians missed diagnosis rates.

With the deep integration of emerging technologies such as Transform, recurrent neural networks (RNN) and diffusion models with traditional models such as CNN and GAN, AI will demonstrate greater auxiliary diagnosis capabilities in the field of medical image processing, and medical image processing involving AI will develop in a smarter, safer and more reliable direction.

References

[1] Y. Shoshan, R. Bakalo, F. Gilboa-Solomon, et al., Artificial Intelligence for Reducing Workload in Breast Cancer Screening with Digital Breast Tomosynthesis, Radiology 2022 Volume 303:1, 69-77.

[2] Q. Han, X. Qian, H. X. Xu, et al., DM-CNN: Dynamic Multi-scale Convolutional Neural Network with uncertainty quantification for medical image classification, Computers in Biology and Medicine, Volume 168, (2024)

[3] H. Ding, N. Huang, Y. Wu and X. Cui, LEGAN: Addressing Intraclass Imbalance in GAN-Based Medical Image Augmentation for Improved Imbalanced Data Classification, in IEEE Transactions on Instrumentation and Measurement, vol. 73, pp. 1-14, 2024, Art no. 2517914.

[4] O. Ronneberger, P. Fischer and T. Brox, U-Net: Convolutional networks for biomedical image segmentation, Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent, pp. 234-241, 2015.

[5] D. W. Dai, C. X. Dong, Q. S. Yan, et al., I2U-Net: A dual-path U-Net with rich information interaction for medical image segmentation, Medical Image Analysis, Volume 97, (2024).

[6] H. N. Veena, K. K. Patil, P. Vanajakshi, et al. An Enhanced RNN-LSTM Model for Fundus Image Classification to Diagnose Glaucoma. SN COMPUT. SCI. 5, 514 (2024).

[7] L. Y. Fu, Y. Z. Chen, W. Ji, F. Yang, SSTrans-Net: Smart Swin Transformer Network for medical image segmentation, Biomedical Signal Processing and Control, Volume 91, (2024).

[8] S. Baba, G. X. Li, and T. Kamiya, Rigid Image Registration for Head MRI Based on 3DCNN Incorporated Global Information. In Proceedings of the 2024 9th International Conference on Biomedical Signal and Image Processing (ICBIP '24). Association for Computing Machinery, New York, NY, USA, 62–65 (2024).

[9] H. Gwon, I. Ahn, Y. Kim, et al., LDP-GAN: Generative adversarial networks with local differential privacy for patient medical records synthesis, Computers in Biology and Medicine, Volume 168, (2024)

[10] J. Chen, J. Mei, X. H. Li, et al., TransUNet: Rethinking the U-Net architecture design for medical image segmentation through the lens of transformers, Medical Image Analysis, Volume 97, (2024)

[11] G. Q. Sun, Y. Z. Pan, W. K. Kong, et al, DA-TransUNet: integrating spatial and channel dual attention with transformer U-net for medical image segmentation, Frontiers in Bioengineering and Biotechnology, Volume 12 (2024)

[12] Q. Zhang, W. Qi, H. Zheng and X. Shen, CU-Net: A U-Net Architecture for Efficient Brain-Tumor Segmentation on BraTS 2019 Dataset, 2024 4th International Conference on Machine Learning and Intelligent Systems Engineering (MLISE), Zhuhai, China, 2024, pp. 255-258.

[13] P. Senapati, A. Basu, M. Deb, et al. Sharp dense U-Net: an enhanced dense U-Net architecture for nucleus segmentation. Int. J. Mach. Learn. & Cyber. 15, 2079–2094 (2024)