ITM Web of Conferences **78**, 02005 (2025) https://doi.org/10.1051/itmconf/20257802005
CSEIT 2025

Optimizing Medical Imaging with AI: Advanced Cnn and Gan Techniques for Enhanced Diagnostic Efficiency

Bowen Ji

Portland College, Nanjing University of Posts and Telecommunications, Nanjing, China

Abstract. In recent years, the explosive growth of medical imaging data and the rapid development of artificial intelligence (AI) have transformed the landscape of diagnostic medicine. This paper investigates the optimization of deep learning algorithms—specifically Convolutional Neural Networks (CNN) and Generative Adversarial Networks (GAN)—to enhance medical image processing. CNNs have proven exceptionally effective in tasks such as image segmentation, lesion detection, and classification by automatically extracting high-level features from raw data. Optimization strategies including parallel computing, quantization, and model pruning have been implemented to accelerate inference speed and reduce energy consumption, making CNNs highly suitable for real-time clinical applications. In contrast, GANs offer unique advantages in generating high-quality synthetic images through adversarial training, thereby addressing issues related to low-resolution imaging and limited data availability. Techniques such as hybrid computing and memory distribution optimization have been explored to boost GAN training efficiency and improve output quality. Despite these advancements, significant challenges remain, including high computational complexity, data privacy concerns, algorithm adaptability, and the necessity for coordinated hardware–software optimization. This paper provides a systematic analysis of the acceleration methods for both CNN and GAN models on AI chips, detailing their strengths, limitations, and practical applications in medical imaging.

1 Introduction

Medical imaging has become indispensable in modern healthcare, serving as a critical tool for diagnosis, treatment planning, and disease monitoring [1,2]. With the advent of high-resolution imaging modalities such as computed tomography (CT), magnetic resonance imaging (MRI), and ultrasound, clinicians now have unprecedented access to detailed anatomical and pathological information. However, this technological progress has also resulted in an exponential increase in the volume and complexity of imaging data. Traditional image processing methods, once sufficient for analyzing simpler data sets, are now challenged by the need for greater accuracy, efficiency, and speed. This has spurred the

p23000519@njupt.edu.cn

integration of artificial intelligence (AI) into medical imaging, where deep learning algorithms have emerged as a transformative force.

Among the deep learning models, Convolutional Neural Networks (CNN) and Generative Adversarial Networks (GAN) have been at the forefront of innovation. CNNs, originally designed for image recognition tasks, have evolved into powerful tools capable of automatically learning hierarchical features from raw medical images. Their ability to perform multi-layered convolutions and deconvolutions enables them to capture intricate spatial details, which is essential for precise image segmentation and lesion detection. Widely used architectures such as U-Net, ResNet, and DenseNet have set new standards in image classification and object detection, facilitating early diagnosis and improving clinical outcomes. The inherent capacity of CNNs to filter and extract salient features reduces the reliance on manual intervention and expert knowledge, thereby streamlining the diagnostic process.

Parallel to the success of CNNs, GANs have introduced a novel approach to medical image analysis [3]. Comprising a generator and a discriminator that engage in adversarial training, GANs are adept at synthesizing high-quality images from low-resolution or incomplete datasets. This capability is particularly valuable in medical imaging, where obtaining high-quality data can be hampered by technical limitations or patient-specific factors. GANs enable not only image enhancement through super-resolution reconstruction and denoising but also facilitate data augmentation by generating realistic synthetic images. Such augmentation is crucial in overcoming the common challenge of limited annotated medical data, which can impede the training of robust diagnostic models.

The integration of AI algorithms with dedicated hardware acceleration further amplifies the impact of these deep learning techniques. Specialized AI chips—such as GPUs, TPUs, FPGAs, and ASICs—are engineered to handle the intensive computations required by CNN and GAN models. Through methods like parallel computing, quantization, and model pruning, these hardware solutions significantly reduce processing times and energy consumption. In clinical environments, where rapid and accurate diagnosis is essential, such improvements can make the difference between timely treatment and diagnostic delay. Moreover, hybrid computing approaches that merge hardware and software optimization techniques are proving instrumental in managing the high computational demands posed by three-dimensional medical imaging data.

Despite the significant advancements, the practical implementation of CNN and GAN algorithms in medical imaging still faces numerous challenges. Data privacy and security remain critical concerns, particularly when dealing with sensitive patient information. The complexity of CNN architectures often requires vast amounts of annotated data, which is not always available due to privacy constraints and the high cost of expert labelling [4,5]. Similarly, while GANs offer remarkable benefits, their training processes are known to be unstable and computationally demanding. Additionally, the integration of these algorithms with AI chips requires a delicate balance between algorithm design and hardware capabilities—a balance that is essential to optimize energy efficiency and ensure reliable performance in real-world applications.

2 Related theories

2.1 Definition and Fundamentals of CNN

Convolutional Neural Network is a deep learning model originally proposed by Yann LeCun. And widely used in the field of computer vision. Convolutional neural networks have become a research hot spot due to their excellent performance in tasks such as image recognition, object detection, and image semantic segmentation, and have been applied in various fields. In a typical CNN, the first few layers are usually alternating between convolutional and down sampling layers, and the last few layers near the output layer are usually fully connected networks. The training process of convolutional neural networks mainly involves learning network parameters such as convolutional kernel parameters and inter layer connection weights of convolutional layers. The prediction process is mainly based on calculating category labels based on input images and network parameters The key to convolutional neural networks is the network structure (including convolutional layers, down sampling layers, fully connected layers, etc.) and back propagation algorithm [2].

The basic principles of CNN acceleration mainly rely on various technical means such as parallel computing, algorithm optimization, hardware optimization, model compression, software optimization, distributed computing, and mixed precision training. Firstly, GPU and TPU can perform parallel computing, which improves efficiency. Secondly, algorithm optimization such as fast convolution algorithms and sparse computing can reduce computational complexity and decrease the number of parameters. In terms of hardware optimization, the use of dedicated hardware such as FPGA and ASIC, as well as memory access optimization, can further improve computing performance. In addition, model compression techniques such as pruning, quantization, and knowledge distillation can reduce computational and storage requirements while maintaining model performance. Software optimization further improves efficiency through the underlying optimization and automatic tuning techniques of deep learning, and reduces the pressure on a single device through distributed computing. Finally, hybrid precision training reduces memory usage and accelerates the computation process. Based on these methods, CNN acceleration technology can significantly improve computational efficiency and is suitable for large-scale model inference.

2.2 Application of CNN in Medical Imaging Acceleration

In recent years, convolutional neural networks have proven their practical value in many fields. Since the birth of AlexNet in 2012, the accuracy of various tasks in computer vision has made a qualitative leap, and some can even exceed that of trained humans [3]. CNN has two main advantages. Firstly, CNN can filter advanced information from raw inputs. Secondly, convolutional neural algorithms and upsampling can greatly reduce the number of samples.

Firstly, there is image segmentation, with traditional methods such as threshold segmentation, region growing, edge detection, and so on [6]. These methods rely heavily on manual labor, have complex structures, and cannot completely eliminate noise. CNN can accurately locate the target area and has many classic structures, such as U-Net (Artificial Intelligence Image Segmentation Algorithm) achieves high-precision segmentation through skip connections between encoder decoder. Next is lesion detection and classification judgment. CNN can perform feature extraction, region localization, and multi-scale detection in lesion detection. Commonly used methods include Faster R-CNN (a classic object detection algorithm that generates candidate regions through Region Proposal Network (RPN) and classifies and regresses each region), YOLO (the detection task is regarded as a regression problem, directly predicting the position and category of lesion regions in the image, with high real-time performance), RetinaNet (which solves the problem of class imbalance through focus loss and is suitable for situations with few lesion regions in medical images), etc. In terms of classification and diagnosis, CNN plays a role in feature learning, end-to-end learning, and multi task learning, simplifying the diagnostic process. Our commonly used methods include ResNet, DenseNet, etc. It has been applied in many aspects, such as in the research on pneumonia infection at the Army Medical University [7].

Five artificial intelligence deep learning models based on chest CT images were established for the classification of bacterial, fungal, common viral, and COVID-19 pneumonia, all of which successfully classified and predicted pneumonia [4]. Finally, there is image reconstruction. CNN can help with super-resolution reconstruction and denoising of medical images, assisting doctors in observing lesions more clearly. CNN can learn the relationship between noisy images and noise free images, such as RED Net, and can better preserve detailed information. In terms of super-resolution reconstruction, SRCNN (Super Resolution Convolutional Neural Network) learns the mapping from low resolution images to high-resolution images through a three-layer convolutional network, using loss functions such as perceptual loss and adversarial loss, and performs multi-scale reconstruction. And CNN can complete data.

2.3 Definition and Role of GAN

Generative Adversarial Networks consist of two networks: the generative network is responsible for generating fake data, and the discriminative network needs to determine whether the input data is true data. GAN requires a training process. This is a game between a generator and a discriminator. The generator generates fake data, and then mixes it with real data to input into the discriminator [8]. The discriminator needs to determine which ones are fake and gradually optimize them. The basic principle of GAN acceleration is to improve computational efficiency, reduce training time and resource consumption through techniques such as parallel computing, model simplification, mixed precision training, gradient optimization, distributed training, pre training and transfer learning, efficient loss function design, hardware acceleration, data augmentation and preprocessing, and progressive training.

2.4 Application of GAN in Medical Imaging Acceleration

GAN has a significant effect on accelerating medical imaging, It can effectively improve image quality, shorten imaging time, and support accurate diagnosis. The various applications of generative adversarial networks in the field of medical imaging can be summarized according to their application purposes, including data augmentation, modal transfer, image segmentation, and image denoising [5].

The diverse medical images generated by GAN can expand the training dataset, improve the generalization ability of deep learning models, and avoid over fitting, Cross modal generation of medical images can also be achieved, This is of great significance for the training and optimization of medical imaging analysis models. The mainstream image conversion models now include Cycle-GAN, The main process of Cycle GAN is to use two generators and two discriminators to achieve mutual transformation between the source domain and the target domain, respectively. Among them, one generator converts an image from the source domain to the target domain, and the other generator converts the converted image from the target domain back to the source domain.

Two discriminators are used to distinguish between the converted and real images. Then, the mapping between the two domains is forcibly modeled as injective through cyclic consistency loss [6]. Image segmentation is the technique and process of dividing an image into several specific regions with unique properties and proposing the target of interest. In the field of medical image processing, accurate segmentation of images is a key step in quantifying clinical parameters and disease classification, which is of great significance for clinical diagnosis, treatment decision-making, and surgical prognosis. GAN performs well in automatic segmentation of medical images, such as tumor or organ segmentation. By generating high-quality segmentation results, GAN reduces the need for manual intervention, improves segmentation accuracy, and accelerates the diagnostic process. In terms of image denoising, it is necessary to reconstruct low-dose, low exposure noisy images into ideal images. Wolterink et al. proposed a generative adversarial network based on voxel loss and adversarial loss for estimating conventional dose CT images from low-dose CT images. Experimental results showed that the network is feasible for denoising low-dose CT images and allows for scoring of coronary artery calcification in low-dose patient CT images with high noise. Combined with compressive sensing technology, GAN can reconstruct high-quality medical images from a small amount of data. This technology significantly shortens imaging time, reduces patient discomfort, while maintaining high image quality.

3 Strengths and limitations

3.1 Advantages of CNN and GAN

CNN networks can directly use raw images as input to automatically learn and extract features, avoiding the complex process of feature definition and parameter setting in traditional algorithms. This makes their performance in various fields better than simply using traditional algorithms, and some can even reach the level of professional technicians [9]. The application of CNN in medical image segmentation has unique advantages. It can extract high-level spatial features of the input image through multi-layer convolution and perform segmentation based on them. In areas that are difficult for the human eye to recognize and distinguish, the effect is particularly significant.

GAN has significant advantages in medical image acceleration processing. Firstly, it can significantly improve image quality, such as restoring high-resolution and clear image details from low-quality data through super-resolution and denoising techniques, providing more reliable basis for clinical diagnosis. Secondly, GAN can shorten imaging time by quickly generating high-quality images from sparse data through image reconstruction and accelerated imaging techniques, improving medical efficiency and reducing patient discomfort. In addition, GAN supports data augmentation and synthesis, which can generate diverse medical images and solve the problem of insufficient data, especially in rare case studies, which is of great significance. Its automation capability performs well in tasks such as image segmentation, registration, and anomaly detection, reducing manual intervention and improving diagnostic efficiency. Finally, GAN supports personalized treatment and multimodal image fusion, generating patient specific images to optimize treatment plans, and integrating medical images from different modalities to provide more comprehensive diagnostic information. These advantages make GAN an important tool in medical image processing, providing strong technical support for precision medicine.

3.2 Deficiencies and Limitations

However, there are still related issues with the application of CNN in this field [10]: (1) high computational complexity, CNN convolution operations and deep network structures require a large amount of resources, especially when processing high pixel images. Currently, many experiments are conducted on the premise of reducing pixels. (2) The memory usage of CNN is too high, especially for deep learning networks that require storing a large number of feature maps and weights, resulting in high memory usage and making it difficult to run on embedded and mobile devices. (3) The risk of over fitting is that CNN models are prone to over fitting, especially in situations where training is insufficient or the model is too complex. These problems are expected to be solved in the short term with the gradual popularization of medical networking. On the other hand, we can also hope for the development of unsupervised learning, which can directly mine features from existing data without relying on human labeling.

Due to the imperfect technology of GAN, there are a series of problems in its application in medical imaging: (1) Most medical images are three-dimensional, and how to improve accuracy and resolution is a problem [11]. (2) How to improve accuracy while ensuring data diversity. (3) How to improve the generalization ability of adversarial networks in medicine, where many models are applied based on exaggerated modalities. These limitations need to be gradually addressed through technological optimization and multi domain cooperation to fully unleash the potential of GAN in medical imaging.

4 Discussion

4.1 Algorithm characteristics and computational requirements

CNN is mainly used in scenarios such as image classification and object detection, with its core being convolution computation [12]. The convolutional layers of CNN involve a large amount of matrix and addition calculations, while AI chips are accelerated through parallel computing and computing cores. GAN consists of a generator and a discriminator. The generator generates images, the discriminator distinguishes images, and GAN performs adversarial optimization through two types of networks. The computational density of GAN is relatively high, and the generator often needs to perform deconvolution calculations, while the discriminator is similar to CNN. AI needs to optimize both the generator and discriminator simultaneously, which requires a large amount of work.

4.2 Energy Efficiency Ratio

The computation of CNN is relatively regular, and energy efficiency can be improved through hardware upgrades. The calculation of GAN is relatively complex, especially the cross training of generator and discriminator, which leads to energy efficiency ratio [13]. AI chips need to optimize energy efficiency through techniques such as dynamic voltage frequency regulation (DVFS).

4.3 Application Scenarios

CNN is widely used in tasks such as lesion detection, image segmentation, and classification in medical imaging. The acceleration of AI chips enables these tasks to be completed in real-time or near real-time. GAN is mainly used for tasks such as data augmentation, image generation, and super-resolution reconstruction in medical imaging. The acceleration of AI chips can improve the quality of generated images and training efficiency.

5 Conclusion

This study has demonstrated that the integration of advanced deep learning algorithms such as Convolutional Neural Networks and Generative Adversarial Networks with dedicated AI hardware can significantly enhance medical imaging applications. The paper systematically analyzed various acceleration techniques, including parallel computing, quantization, model pruning, and hybrid computing, to optimize the performance of CNNs and GANs on AI chips. CNNs, with their inherent ability to extract high-level features from raw images, have proven to be highly effective in tasks like image segmentation, lesion detection, and classification. The adoption of architectures like U-Net, ResNet, and DenseNet has revolutionized the precision and efficiency of these diagnostic processes, reducing reliance on manual intervention.

Similarly, GANs have shown tremendous potential in improving image quality by generating high-resolution images and facilitating data augmentation through adversarial training. Techniques to enhance GAN training efficiency, such as memory distribution optimization and dynamic voltage frequency regulation, were also discussed, highlighting the importance of balancing computational demands with energy efficiency. Despite these advancements, challenges persist, including the high computational complexity of deep models, substantial memory usage, and issues of data privacy and algorithm adaptability. These challenges underline the necessity for ongoing improvements in both algorithmic design and hardware integration, ensuring that the benefits of AI-driven medical imaging are fully realized in clinical settings.

References

1. Tian, M., Lin, L., Zhang, B., Wu, S.: 'College of Life Science and Bioengineering', Beijing University of Technology, 2023, 100124, (Beijing).
2. Chang, L., Deng, X., Zhou, M., Wu, Z., Yuan, Y., Yang, S., Wang, H.: 'Convolutional neural networks in image understanding', Journal of Automation, 2016, 42, (9), pp. 1300–1312.
3. Krizhevsky, A., Sutskever, I., Hinton, G.: 'ImageNet classification with deep convolutional neural networks', Commun. ACM, 2017, 60, (6), pp. 84–90.
4. Zeng, M., Zhao, N., Wang, X., Pang, H., Hou, W., Liu, L., Wang, X., Gong, J., Zeng, W., Cui, H., Chen, W.: 'Establishment of a chest-CT-image-based classification deep learning model for pneumonia prediction', Journal of Army Medical University, 2023, 45, (21), pp. 2266–2274.
5. Zhu, X., Liu, Z., Cambria, E., Yu, X., Fan, X., Chen, H., Wang, R.: 'A client–server based recognition system: Non-contact single/multiple emotional and behavioral state assessment methods', Comput. Methods Programs Biomed., 2025, 260, pp. 108564.
6. Wang, R., Zhu, J., Wang, S., Wang, T., Huang, J., Zhu, X.: 'Multi-modal emotion recognition using tensor decomposition fusion and self-supervised multi-tasking', Int. J. Multimed. Inf. Retr., 2024, 13, (4), pp. 39.
7. Wang, F., Ju, M., Zhu, X., Zhu, Q., Wang, H., Qian, C., Wang, R.: 'A geometric algebra-enhanced network for skin lesion detection with diagnostic prior', J. Supercomput., 2025, 81, (1), pp. 1–24.
8. Zhao, Z., Zhu, X., Wei, X., Wang, X., Zuo, J.: 'Application of workflow technology in the integrated management platform of smart park', IEEE Adv. Inf. Manag., Communicates, Electron. Automat. Control Conf., 2021, 4, pp. 1433–1437.
9. Zhang, Y., Zhao, H., Zhu, X., Zhao, Z., Zuo, J.: 'Strain measurement quantization technology based on DAS system', IEEE Adv. Inf. Manag., Communicates, Electron. Automat. Control Conf., 2019, pp. 214–218.
10. Li, H., Zhou, W., Chen, Y., Zhang, M.: 'Domain independent Cycle-GAN for cross modal medical image generation', Frontiers of Data & Computing, 2024, 6, (2), pp. 80–88.
11. Yang, P., Zhou, Y., Xing, G., Tian, Z., Xu, X.: 'Applications of convolutional neural network in biomedical image', 2021.

12. Zheng J, Yang C, Zhang T, et al. Dynamic Spectral Graph Anomaly Detection[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2025, 39(12): 13410-13418.
13. Gao M, Sun J, Li Q, et al. Towards trustworthy image super-resolution via symmetrical and recursive artificial neural network[J]. Image and Vision Computing, 2025: 105519.