AI and Data Science Journal
ISSN: 3078-7394
DOI: https://doi.org/10.61784/adsj3019

CURRENT ADVANCED MEDICAL IMAGE PROCESSING METHOD—INTEGRATION AND INNOVATION OF DEEP LEARNING MODELS AND TRADITIONAL ALGORITHMS

ShouTong Huang
Electrical and Electronics Engineering Department, Ningxia University, Yinchuan 750021, Ningxia, China.
Corresponding Email: 13752673836@139.com

Abstract: Medical image processing has witnessed remarkable progress in recent years, driven by the integration of advanced algorithms and artificial intelligence techniques. This review comprehensively examines the state-of-the-art methods in medical image processing, both domestically and internationally, with a focus on deep learning models and other prominent algorithms. We delve into their applications across various medical imaging modalities, analyze their strengths and limitations, and discuss future development trends, aiming to provide valuable insights for researchers and practitioners in this field.
Keywords: Algorithms; Deep learning; Image processing; Medical image segmentation

1 INTRODUCTION

Medical imaging technologies like computed tomography (CT), magnetic resonance imaging (MRI), ultrasound imaging, and positron emission tomography (PET) have become indispensable tools in modern healthcare for disease diagnosis, treatment planning, and monitoring. However, the vast amount of complex image data generated poses significant challenges in interpretation and analysis. Medical image processing, which involves the use of computational algorithms to enhance, analyze, and interpret these images, has therefore emerged as a crucial interdisciplinary field. Medical image processing is a cornerstone technology for disease diagnosis, surgical planning, and treatment efficacy assessment. Recent years have seen an explosion of innovative methods, particularly, deep learning (DL)-based approaches have emerged as the dominant approach in medical image analysis due to their strengths in automated feature extraction and complex pattern recognition, that have transformed the landscape of medical image processing. However, traditional algorithms (e.g., fuzzy clustering, graph-cut methods) remain indispensable in specific scenarios. This paper aims to provide an in-depth review systematically of the latest advancements in medical image segmentation, registration, classification, keypoint detection, and generative models, while addressing challenges such as multimodal fusion and few-shot learning.

2 OVERVIEW OF MEDICAL IMAGE PROCESSING

Medical image processing typically involves several key steps:

2.1 Image Acquisition

Different imaging modalities capture physiological information in distinct ways. CT uses X-rays to generate cross-sectional images, MRI relies on magnetic fields and radio waves to produce detailed soft tissue contrast, ultrasound employs sound waves for real-time imaging, and PET detects gamma rays from radiotracers to map metabolic activity. Each modality has its unique advantages and limitations, influencing subsequent processing steps.

2.2 Image Preprocessing

This stage aims to improve image quality and prepare data for further analysis. Common techniques include:
Filtering: To reduce noise while preserving important anatomical structures. For example, Gaussian filters smooth out random noise, while median filters are effective against salt-and-pepper noise.
Intensity Normalization: Adjusting pixel intensity values to a standard range to eliminate variations caused by different imaging devices or acquisition settings.
Artifact Reduction: Removing unwanted artifacts like metal artifacts in CT or motion artifacts in MRI that can distort image content.

2.3 Image Analysis and Interpretation

This encompasses various tasks such as: Quantitative Measurement: Extracting numerical features like size, shape, and texture of anatomical structures for objective assessment.

2.4 Image Storage and Transmission

Efficient storage and secure transmission of medical images are essential for healthcare institutions. This involves compression techniques to reduce data size without significant loss of diagnostic information and standardized formats like DICOM (Digital Imaging and Communications in Medicine) for interoperability.

2.5 Image Segmentation

Identifying and delineating regions of interest (e.g., tumors, organs) from the background.
Deep Learning Methods: U-Net and Its Variants: U-Net, with its skip connections that integrate shallow-layer details and deep semantic information, has become the benchmark model for medical segmentation. U-Net++ further optimizes gradient propagation through nested dense connections, achieving a 12% accuracy improvement in liver tumor segmentation[1]. DeepLab Series: By leveraging dilated convolutions to expand the receptive field, DeepLab addresses boundary ambiguity in lung CT nodule segmentation[2]. Generative Adversarial Networks (GANs):
Adversarial training enables high-quality mask generation. For example, CycleGAN has demonstrated remarkable performance in cross-modal segmentation between MRI and CT images[3].
Traditional Methods: Fuzzy C-Means Clustering (FCM): This method handles low-contrast images via pixel fuzzification, outperforming traditional thresholding in 3D bone reconstruction. Region Growing: Combined with morphological priors, it is widely used for preliminary brain tumor localization.
Challenges: Limited generalization under small-sample scenarios necessitates integration with transfer learning (e.g., ResNet pre-trained models).

2.6 Medical Image Registration

Aligning images from different modalities or time points to a common coordinate system for comparative analysis.
Supervised and Unsupervised Learning: Supervised Methods: These rely on ground-truth annotations to train registration networks using simulated deformation fields, but suffer from high annotation costs. Unsupervised Methods: Based on image similarity loss (e.g., mutual information) and deformation regularization, they achieve accuracy comparable to traditional optimization algorithms in brain multimodal registration.
Cross-Modal Registration: Feature-Level Alignment: Deep Belief Networks (DBNs) extract modality-invariant features, reducing errors by 20% in PET-MRI registration. *G. Medical Image Classification and Detection*
Categorizing images or image regions into predefined classes (e.g., normal vs. pathological).
Classification Models: ResNet and Attention Mechanisms: Residual connections mitigate gradient vanishing, while channel attention modules (e.g., SE blocks) achieve an AUC of 0.96 in breast cancer X-ray classification[4]. Few-Shot Learning: Meta-learning enables high-precision training with only 50 samples for thyroid ultrasound image classification.
Keypoint Detection: Regression-Classification Hybrids: Dual-path networks simultaneously predict keypoint coordinates and class probabilities, achieving < 2mm error in spinal anatomical landmark detection.

3 DEEP LEARNING MODELS IN MEDICAL IMAGE PROCESSING

Deep learning, a subset of machine learning based on artificial neural networks with multiple layers, has revolutionized medical image processing due to its powerful feature learning and representation capabilities.

3.1 Convolutional Neural Networks (CNNs)

CNNs are the backbone of many medical image analysis tasks. Their architecture consists of convolutional layers that automatically extract hierarchical features from image data.
Applications in Segmentation: U-Net, a popular CNN architecture, has achieved remarkable success in medical image segmentation. It employs a contracting path to capture context and an expanding path for precise localization, making it highly effective for tasks like tumor segmentation in brain MRI or lung segmentation in CT scans. Variants like Attention U-Net incorporate attention mechanisms to focus on relevant regions, improving segmentation accuracy.
Applications in Classification: For disease classification, CNNs can be trained on large datasets to distinguish between different pathologies. For instance, in skin cancer detection, CNNs analyze dermoscopic images to differentiate malignant melanomas from benign lesions with high sensitivity and specificity. Transfer learning, where pre-trained models on natural image datasets (e.g., ImageNet) are fine-tuned on medical images, has proven effective given the limited availability of annotated medical data[5].
Applications in Registration: Some recent approaches utilize CNNs to learn registration parameters directly from image pairs, offering faster and more accurate alignment compared to traditional intensity-based registration methods.

3.2 Recurrent Neural Networks (RNNs)

RNNs are designed to handle sequential data, making them suitable for processing time-series medical images or images with spatial dependencies.

Applications in Dynamic Imaging: In cardiac MRI or PET scans that capture physiological changes over time, RNNs model the temporal dynamics to analyze heart function or metabolic processes. For example, they can track the contraction and relaxation of heart chambers across cardiac cycles.
Applications in Image Captioning: Generating textual descriptions of medical images, RNNs combined with CNNs first extract visual features and then generate coherent sentences explaining findings like "a well-circumscribed mass in the upper lobe of the right lung."

3.3 Generative Adversarial Networks (GANs) : Generative Models and Data Augmentation in Image Synthesis and Reconstruction

GANs consist of a generator and a discriminator network that compete with each other, leading to the generation of realistic synthetic images.
Applications in Data Augmentation: Medical image datasets are often limited in size and diversity. GANs can generate synthetic but realistic-looking images to augment training data, improving the robustness of models trained on small datasets. For instance, in rare disease diagnosis, GANs create additional pathological cases to balance class distribution.
Applications in Super-Resolution Reconstruction: Enhancing the resolution of low-quality medical images, GANs learn to map low-resolution inputs to high-resolution outputs, recovering finer details that might be critical for diagnosis. This is particularly useful in ultrasound imaging where resolution is inherently limited. DL models based on convolutional sparse coding improve resolution by 30% in low-dose CT images.
Applications in Synthetic: Pix2Pix generates synthetic MRI images to address data scarcity, while StyleGAN2 outperforms traditional methods in pathological slice generation[6].

3.4 3D Convolutional Neural Networks

Extending CNNs to three dimensions, 3D CNNs process volumetric medical images directly, capturing spatial relationships in all dimensions.
Applications in Volumetric Analysis: For brain MRI or CT volumes, 3D CNNs analyze the entire 3D structure to detect abnormalities like brain tumors or hemorrhages. They can also perform organ segmentation in abdominal CT scans, considering the full 3D context for more accurate delineation compared to 2D slice-wise approaches[7].

4 COMPARISON AND DISCUSSION OF DIFFERENT METHODS

4.1 Strengths and Limitations

Deep Learning Models: Their major advantage lies in automated feature learning from large data, achieving high accuracy in complex tasks. However, they require substantial annotated training data, which is often difficult and time-consuming to obtain in the medical field. They are also computationally intensive and lack interpretability, making it hard for clinicians to trust and understand their decisions.
Traditional Algorithms: Generally more interpretable and less data-hungry. For example, filter-based methods have clear mathematical formulations, and thresholding relies on intuitive intensity differences. But they often rely on handcrafted features and may struggle with complex variations in medical images.

4.2 Integration Approaches

Combining deep learning with traditional algorithms can leverage their respective strengths. For instance, using traditional filtering as a preprocessing step to enhance image quality before feeding into a deep learning model, or employing clustering to initialize parameters for deep learning segmentation networks[8].

5 CHALLENGES AND FUTURE DEVELOPMENT TRENDS

5.1 Data and Model Limitations

Annotation Costs: Semi-supervised learning (e.g., Mean Teacher) and weak-label methods (e.g., image-level labels) are promising solutions.
Model Interpretability: Visualization tools like Grad-CAM provide lesion localization evidence in lung cancer detection.

5.2 Multimodal and Interdisciplinary Integration

Multi-Task Learning: End-to-end frameworks (e.g., V-Net) that jointly handle segmentation, classification, and registration reduce processing time by 50% in orthopedic surgical planning. Integration of Traditional and Deep Learning: Cascade models combining fuzzy clustering and U-Net balance efficiency and accuracy in cytomorphological analysis.

5.3 Multi-Modal Fusion

Integrating information from multiple imaging modalities (e.g., combining CT and PET) within deep learning frameworks to capture complementary information for more comprehensive analysis.

5.4 Explainable AI

Developing deep learning models with inherent interpretability or creating techniques to explain their decision-making processes, crucial for clinical acceptance.

5.5 Real-Time Processing

Optimizing algorithms for speed to enable real-time medical image analysis during surgeries or emergencies, possibly through model compression and hardware acceleration.

5.6 Personalized Medicine

Tailoring image analysis to individual patients by incorporating their unique genetic, clinical, and imaging data into models.

5.7 How to Solve the Problem of Small-Shot Learning in Medical Image Processing

In medical image processing, small-shot learning is an important research direction, because the acquisition cost of medical data is high and the annotation is difficult, resulting in a limited amount of training data available. To address this issue, researchers have proposed a variety of methods and techniques. Here are some of the main solutions:

Generative Adversarial Network (GAN) is a method of generating high-quality data through adversarial training of generators and discriminators. In medical image processing, GAN can be used to generate more training samples, thereby improving the generalization ability of the model. For example, by generating artifact images that resemble real data, you can increase the diversity of the training data, thereby improving the robustness and accuracy of the model.

Meta-learning is a method of adapting quickly to new tasks by learning how to do so. In medical image processing, meta-learning can be used to quickly adjust model parameters with a small number of samples, thereby improving the adaptability and performance of the model. For example, with meta-learning, the model can be fine-tuned on a small number of samples to achieve better classification results.

Graph Neural Network (GNN) is a deep learning method based on graph structured data. In medical image processing, GNN can be used to extract local features in images and disseminate information through graph structures, so as to improve the classification performance of models. For example, by constructing a graph neural network model, the edge contours and tumor details in the breast ultrasound images can be effectively extracted, so as to improve the accuracy of classification.

Contrastive learning is a method of learning feature representations by comparing pairs of positive and negative samples. In medical image processing, contrastive learning can be used to enhance the model's ability to distinguish between different types of samples. For example, through contrastive learning, the model can learn more robust representations of features, thereby improving the accuracy and robustness of classification.

Multi-scale feature extraction is a method of extracting image features through convolutional kernels of different scales. In medical image processing, multi-scale feature extraction can be used to capture feature information at different scales, thereby improving the performance of the model. For example, with multi-scale design, the model can learn effective target feature information from multiple perspectives at the same stage.

Attribute-based small-shot classification algorithm is a method to improve classification accuracy by using image attribute information. In medical image processing, attribute-based small-shot classification algorithms can be used to make up for the shortcomings of traditional metric learning algorithms in the case of insufficient data. For example, by introducing attribute distribution similarity, the accuracy and robustness of classification can be effectively improved.

Superpixel and pseudo-labeling is a way to reduce the need for annotation by generating superpixel tags. In medical image processing, superpixels and pseudo-labels can be used to reduce the workload of manual annotation, thereby improving the training efficiency of models. For example, by using hyperpixels and corresponding pseudo-labels, an unsupervised learning method can be implemented, which improves the generalization ability of the model.

Self-supervised pre-training is a method of pre-training a model through self-supervised tasks. In medical image processing, self-supervised pre-training can be used to improve the initial performance of the model, thereby reducing the dependence on large amounts of annotated data. For example, with self-supervised pre-training, the model can be fine-tuned on a small number of samples to achieve better classification results.

These methods and technologies have a wide range of application prospects in different application scenarios.

5.8 What are the Latest Use Cases of Generative Adversarial Networks (GANs) in Medical Image Generation and Augmentation

The latest use cases of Generative Adversarial Networks (GANs) in medical image generation and enhancement include the following:

Medical Image Compositing: Medfusion: Medfusion is a Conditional Latent Variable Diffusion Model (Latent DDPM) designed for medical image generation. It was trained and evaluated on three datasets, AIROGS, CheXpert, and CRCDX, and compared with GANs. The results showed that Medfusion surpassed GANs in terms of diversity (Recall), achieving scores of 0.40, 0.36 and 0.42 on AIROGS, CRMDX and CheXpert datasets, respectively, while GANs scored 0.19, 0.02 and 0.17, respectively. In addition, Medfusion achieved fidelity (Precision) comparable to or higher than GANs on all three datasets. This suggests that Medfusion is a GANs-based model, but with higher performance.
Medical Image Enhancement: GANs are widely used in the enhancement of medical images to improve the performance of CNNs in tasks such as liver lesion classification. For example, GANs can be used to generate new medical images, such as those of the brain, spine, and other organs, to help improve the accuracy and efficiency of diagnosis.
Multimodal Image Compositing: GANs also have important applications in multimodal image synthesis. For example, 3D Egenerative Generative Adversarial Networks (E-GAN) is used for PET to T1-weighted MRI translation, improving the quality and efficiency of medical image processing by generating high-quality 3D images.
Medical Image Segmentation and Classification: GANs are also widely used in medical image segmentation and classification. For example, GANs can be used for tasks such as lung nodule detection, dementia diagnosis, lung dataset classification, breast cancer detection, lung cancer image enhancement, COVID-19 screening, brain environment data enhancement, pneumonia and COVID-19 detection, Alzheimer's disease classification, and more.
Unsupervised Representation Learning: GANs also have important applications in unsupervised representation learning. For example, deep convolutional generative adversarial networks (DCGANs) demonstrate their potential in unsupervised learning by learning hierarchical representations from object parts to scenes.
Image Data Enhancement: GANs are also widely used in image data enhancement. For example, GANs can be used to generate new image data to increase the diversity and number of training datasets, thereby improving the generalization ability of the model.
GANs are widely used in medical image generation and enhancement, covering many aspects from image synthesis and enhancement to multimodal image synthesis, segmentation and classification[9].

5.9 What are the New Developments of Multimodal Data Fusion Technology in Medical Image Processing

According to the information I searched, the multimodal data fusion technology in medical image processing has made significant progress in recent years. Here are some of the new developments and trends: Application of Deep Learning in Multimodal Medical Image Fusion: Deep convolutional neural networks (CNNs) have played an important role in multimodal medical image fusion. Through ensemble learning methods, these networks are able to effectively overcome the limitations of a single modality, such as noise, artifacts, or incomplete information, to provide a more comprehensive and information-rich representation of images. The multimodal medical image fusion helps radiologists and clinicians make accurate diagnosis, treatment planning, and patient monitoring. Diversity and Innovation of Multimodal Data Fusion Technology: Its research covers deep learning, graph neural networks, Siamese networks, memory attention mechanisms, converter networks, multimodal semantic image segmentation, multimodal information extraction, multimodal data fusion, multimodal image classification, multimodal weight sharing, feature fusion, brain tumor segmentation, multimodal heterogeneous data learning, GAN technology of multimodal data fusion, and multimodal medical image fusion technology based on NSCTD and DTCWT. The diversity and innovation of these technologies open up new possibilities for improving diagnostic accuracy, treatment outcomes, and patient care.
Comparison of Multimodal Data Fusion Methods: By comparing the three techniques of delayed fusion, early fusion, and sketch representation, the study shows that choosing the appropriate fusion technique is crucial for constructing multimodal representations. Experimental results show that these techniques can significantly improve the performance of classification tasks.
Amazon Reviews, MovieLens25M, and MovieLens1M datasets were used in the experiment, covering three modalities: text, image, and graphic data. Advances in clinical application of multimodal data fusion: The increasing use of AI technology in fusing electronic health records (EHRs) with medical imaging data is critical to enabling precision medicine. In particular, advances in machine learning (ML) across different data modalities provide multimodal insights for clinical applications. The study found that the number of studies fusing image data with EHR increased significantly between 2020 and 2021, indicating the increasing importance of multimodal data fusion in clinical applications.
Challenges and Future Directions of Multimodal Data Fusion: Despite the positive results, multimodal medical image fusion still faces some challenges, such as selecting the right ensemble learning technology, optimizing the network architecture and strategy, and the availability of large-scale annotated datasets and computing resources.
Future research needs to further explore these challenges to advance the development of multimodal medical image fusion technology.
In summary, the multimodal data fusion technology in medical image processing has made remarkable progress in deep learning, algorithm innovation and clinical application[10].

6 CONCLUSION AND OUTLOOK

The field of medical image processing is rapidly evolving with continuous breakthroughs in algorithm development. Deep learning has propelled medical image processing into a new era, deep learning models have shown great potential but still face challenges in model robustness, computational efficiency, clinical applicability, data availability, computational resources, and interpretability. Traditional algorithms remain relevant and can be synergistically combined with advanced methods. Looking ahead, interdisciplinary collaboration between computer scientists, medical imaging experts, and clinicians will be essential to translate these innovative techniques into clinical practice, ultimately improving patient care and outcomes. Future research should focus on multimodal data fusion, lightweight model design (e.g., MobileNet variants), and synergistic innovation with traditional algorithms to accelerate clinical translation.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015: 326-334.
[2] Chen L C, Papandreou G, Kokkinos I, et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2018, 40(4): 834-848.
[3] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS), 2014: 27.
[4] Zong Y, Li D, Xia S, et al. Deep Learning in Medical Image Analysis. Annual Review of Biomedical Engineering, 2019, 21: 203-227.
[5] Krizhevsky A, Sutskever I, Hinton G E. ImageNet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems (NeurIPS), 2015, 25.
[6] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770-778.
[7] Vaswani A, Shazeer N, Parmar N, et al. Attention Is All You Need. Advances in Neural Information Processing Systems (NeurIPS), 2017, 30.
[8] Kingma D P, Ba J. Adam: A Method for Stochastic Optimization. International Conference on Learning Representations (ICLR), 2015.
[9] LeCun Y, Bengio Y, Hinton G. Deep Learning. Nature, 2015, 521(7553): 436-444.
[10] Esteva A, Robicquet A, Ramsundar B, et al. A Guide to Deep Learning in Healthcare. Nature Medicine, 2017, 25(1): 24-26.