Journal of Artificial Intelligence Practice (2025)
Clausius Scientific Press, Canada
DOI: 10.23977/jaip.2025.080319
ISSN 2371-8412 Vol. 8 Num. 3

A Review of the Basic Applications of Machine Vision in Medical Image Segmentation

Xueju Hao

School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, China
1450946940@qq.com

Keywords: Machine Vision; Medical Image Segmentation; CT; MRI; Clinical Auxiliary Diagnosis

Abstract: Medical image segmentation is a core link in clinical diagnosis, treatment planning, and efficacy evaluation, and its accuracy directly affects the scientificity of medical decisions. With the rapid development of machine vision technology, medical image segmentation based on Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) has become a research hotspot in the field of medical artificial intelligence. This paper focuses on two core scenarios: organ segmentation (e.g., liver, kidney) and lesion segmentation (e.g., tumor), systematically reviewing the basic applications and clinical value of machine vision segmentation technology. First, it combs the development history and core methods of segmentation technology, then introduces the characteristics and application scenarios of classic datasets such as BraTS and LiTS, deeply analyzes key issues currently facing the field including scarcity of annotated data and inconsistent image formats across different hospitals, and discusses the preliminary integration scenarios of technology with clinical diagnosis. Research shows that machine vision segmentation technology can significantly improve the efficiency and accuracy of medical image analysis, providing objective and quantitative reference for clinical practice. However, continuous breakthroughs are still needed in data standardization, model generalization, and clinical adaptability.

1. Introduction

The innovation of medical imaging technology has promoted the transformation of clinical diagnosis models. As representatives of non-invasive imaging technologies, CT and MRI have been widely used in the screening and diagnosis of diseases in various organs of the body. As a key step to extract target regions (organs, lesions, etc.) from original images, image segmentation is a prerequisite for lesion localization, volume measurement, and treatment plan formulation. Traditional manual segmentation relies on physicians' professional experience, which has limitations such as strong subjectivity, time-consuming and labor-intensive processes, and large individual differences. Especially when processing large-scale image data or small lesions, it is prone to missed diagnosis, misdiagnosis, or segmentation errors.

Machine vision technology, with its automated and high-precision feature extraction and pattern recognition capabilities, provides a new solution for medical image segmentation. By simulating the human visual system and combining image processing, machine learning, and deep learning methods, this technology realizes automatic or semi-automatic segmentation of target regions in medical images. In recent years, deep learning-based segmentation models have shown excellent performance in CT and MRI image processing, achieving breakthrough progress in organ segmentation (such as liver and kidney) and lesion segmentation (such as tumor) tasks.

This paper focuses on the basic applications of machine vision in medical image segmentation, taking CT and MRI images as the core research objects, and conducts a review around two major clinical needs: organ segmentation and lesion segmentation. First, it elaborates on the core theoretical basis of segmentation technology, then analyzes typical application scenarios and values in detail; next, it introduces classic datasets and evaluation systems supporting technological development; further discusses the main challenges facing the current field; finally, this paper summarizes the integration status of technology with clinical diagnosis and looks forward to future development trends, providing reference for related research and clinical applications.

2. Fundamentals of Machine Vision Segmentation Technology

2.1 Technological Development History

The application of machine vision in medical image segmentation has two stages: traditional and deep learning methods. Traditional methods focus on manually designed features, including threshold segmentation, edge detection, and region growing. Threshold segmentation separates targets from the background by setting grayscale thresholds, which is simple and computationally efficient but has poor adaptability to uneven - grayscale images and struggles with complex anatomical structures. Edge detection forms target boundaries by extracting pixels with sudden grayscale changes, but its results are easily disturbed by noise. Region growing expands regions based on pixel similarity, retaining spatial continuity well but being sensitive to seed - point selection and prone to over - or under - segmentation.

Deep learning methods have changed the technical path of medical image segmentation. Their core advantage is automatically learning deep semantic features without manual feature engineering. Convolutional Neural Networks (CNNs), as the core model, capture local features through convolution, combine pooling and deconvolution for feature fusion and resolution recovery[7], and are suitable for 2D and 3D segmentation. The U - Net model, a classic architecture, uses an encoder - decoder symmetric structure and skip connections to solve feature loss and positioning accuracy problems, becoming the benchmark for organ and lesion segmentation[1]. Subsequently, researchers proposed improved models like U - Net++, Attention U - Net, and 3D U - Net[4].

In recent years, Transformer models, successful in natural language processing with self - attention mechanisms, are widely used in medical image segmentation. They can capture long - range dependencies and model global features. Hybrid models with CNNs (e.g., Swin Transformer U - Net) perform better in complex anatomical structure segmentation[6]. Moreover, weakly supervised, semi - supervised, and unsupervised learning methods are research hotspots to address the scarcity of annotated data and lower the technological implementation threshold.

2.2 Core Technical Characteristics

Machine vision segmentation technology has three core characteristics in medical image applications. First, it has a high degree of automation, which can significantly reduce manual intervention and shorten segmentation time. Traditional manual segmentation of a 3D CT image takes several hours, while deep learning-based segmentation models can complete the processing in minutes, greatly improving clinical work efficiency. Second, it has high segmentation accuracy. The features learned by the model through large-scale data training have stronger generalization ability, which can accurately identify the fine structure of target regions, and the segmentation results are highly consistent with the gold standard annotated by medical experts. Third, it has strong quantitative analysis capabilities. After segmentation, parameters such as volume, surface area, and grayscale mean of target regions can be automatically calculated, providing objective data support for clinical diagnosis and efficacy evaluation, and avoiding subjective errors in manual measurement.

3. Core Application Scenarios and Clinical Value

3.1 Organ Segmentation Applications

Organ segmentation is a basic medical image analysis task. Its goal is to extract specific organ anatomical structures from CT or MRI images, providing a basis for disease diagnosis, surgical planning, and organ function evaluation. The liver and kidneys, important metabolic and excretory organs, have extensive clinical applications of segmentation technology.

Liver segmentation is based on CT and MRI images, each with advantages. CT images have high spatial resolution and clear tissue density contrast, suitable for liver tumor screening, cirrhosis diagnosis, and surgical planning. Core problems in CT - based liver segmentation include grayscale similarity with surrounding tissues, blurred edges, and lesion interference. Deep learning models can effectively distinguish the liver from surroundings.

MRI images offer multi - sequence imaging, providing abundant soft - tissue contrast information, suitable for liver lesion qualitative diagnosis and functional evaluation. Challenges in MRI - based liver segmentation are image artifacts, uneven grayscale, and information fusion. Multi - modal fusion segmentation models improve robustness.

Kidney segmentation clinical applications focus on kidney disease diagnosis, transplantation evaluation, and urinary system surgical planning. CT images show kidney position, size, and shape, suitable for kidney stone and cancer screening. CT - based kidney segmentation needs to handle adhesion with surrounding tissues, and difficulty increases with hydronephrosis or tumors. MRI images are better for evaluating kidney functional status, providing a basis for chronic kidney disease monitoring.

The clinical value of organ segmentation technology is reflected in three aspects. Firstly, it assists surgical planning, like in liver transplantation to calculate donor - recipient matching and reduce surgical risk. Secondly, it evaluates organ function by monitoring volume changes combined with imaging parameters. Thirdly, it supports multi - center research by providing a unified data basis for large - scale clinical studies.

3.2 Lesion Segmentation Applications

Lesion segmentation is a core clinical need in medical image segmentation, aiming to extract lesion regions from organ images for qualitative analysis, staging, treatment planning, and efficacy evaluation. Tumor segmentation, as the main task, is significant in clinical diagnosis and treatment of diseases like brain and liver tumors.

Brain tumor segmentation is based on MRI images. The BraTS dataset has advanced this field[2]. Brain tumors' irregular shape, blurred boundaries, and invasion of surrounding tissue make traditional methods ineffective. Deep - learning models learn multi - modal MRI features to accurately segment the tumor core, edema, and whole tumor regions. For example, 3D U - Net and its variants can capture 3D tumor structure, with segmentation accuracy approaching medical experts' level.

The clinical value of brain tumor segmentation is shown in multiple links. In diagnosis, it helps determine tumor location, size, and invasion range. In treatment planning, it's the basis for radiotherapy target delineation. In efficacy evaluation, it allows quick judgment of treatment effect by comparing tumor volume changes.

Liver tumor segmentation uses CT and MRI images, and the LiTS dataset supports research. Liver tumors have complex imaging manifestations, and different types vary greatly in grayscale and shape. Some small tumors with blurred boundaries increase segmentation difficulty. CT - based segmentation needs to solve grayscale confusion between blood vessels and tumors, while MRI multi - sequence imaging improves accuracy[8].

Clinical applications of liver tumor segmentation include treatment plan formulation and efficacy monitoring. For resectable tumors, it helps plan resection range. For inoperable advanced liver cancer, it enables accurate formulation of interventional and radiotherapy plans. Regular monitoring of tumor volume and shape helps evaluate treatment effect and optimize strategies.

Besides brain and liver tumors, machine vision segmentation technology is used for various lesions like pulmonary nodules, kidney cancer, and breast cancer. Different lesions have different imaging features and clinical needs, but their core value is to provide accurate lesion information for clinical practice, assist doctors in making decisions, and improve diagnosis and treatment efficiency and effects.

4. Datasets and Evaluation Metrics

4.1 Classic Datasets

Annotated data is the basis for the training and verification of machine vision segmentation models. High-quality public datasets can promote technical exchange and standardized evaluation. Several classic datasets have been formed in the field of medical image segmentation, among which datasets such as BraTS and LiTS are most widely used in organ and lesion segmentation research.

The BraTS dataset, fully known as the Brain Tumor Segmentation Challenge dataset, is released by the Medical Image Computing and Computer-Assisted Intervention Society (MICCAI), focusing on MRI image segmentation tasks of brain tumors. The dataset includes multi-center, multi-modal brain tumor MRI images, covering four sequences: T1, T1-enhanced, T2, and FLAIR. Each image is annotated by multiple medical experts for the tumor core region, edema region, and whole tumor region. The scale of the dataset has expanded year by year, and the latest version includes imaging data of hundreds of patients, providing a unified standard for the training, verification, and comparison of brain tumor segmentation models[5]. The promotion of the BraTS dataset has greatly promoted the development of brain tumor segmentation technology and become the core benchmark for technical evaluation in this field.

The LiTS dataset, namely the Liver and Tumor Segmentation Benchmark dataset, is also released by MICCAI, focusing on CT image segmentation of the liver and liver tumors[3]. The dataset includes abdominal CT images from different hospitals, and each data is annotated with the liver contour and liver tumor region, covering tumor cases of different sizes, shapes, and positions. The characteristic of the LiTS dataset is strong data diversity, which can reflect various complex situations in clinical practice, such as abnormal liver shape, multiple tumors, and adhesion between blood vessels and tumors, providing an important basis for evaluating the generalization ability of models[3].

In addition to the above datasets, there are several commonly used public datasets in the field of medical image segmentation. The CHAOS dataset focuses on CT and MRI image segmentation of multiple abdominal organs (such as the liver, kidneys, and spleen), providing multi-modal and multi-organ annotated data. The KiTS dataset focuses on CT image segmentation of the kidneys and kidney tumors, including detailed annotations of the kidneys, tumors, and renal pelvis. The MSD dataset covers more than 10 types of organ and lesion segmentation tasks, providing a unified evaluation framework. The common value of these datasets is to provide researchers with standardized training and verification data, avoiding the uneven technological development caused by data barriers, and promoting the rapid iteration of segmentation technology.

4.2 Evaluation Metrics

The evaluation of medical image segmentation results should combine clinical needs and technical characteristics. Commonly used evaluation metrics are divided into three categories: similarity, distance, and volume metrics, which quantify the consistency between segmentation results and the gold - standard annotated by medical experts.

The Dice Similarity Coefficient (DSC), the most common similarity metric, calculates the ratio of the intersection of segmentation results and the gold - standard to the union. Ranging from 0 to 1, a value closer to 1 means higher accuracy. It comprehensively reflects accuracy and completeness, is sensitive to small lesion segmentation, and is the core metric for organ and lesion segmentation tasks.

The Intersection over Union (IoU), or Jaccard coefficient, calculates the ratio of the intersection to the entire set. It is highly correlated with DSC, evaluates the overlap of segmentation regions, is easy to calculate in engineering, and is often used in loss - function design during model training.

The Hausdorff Distance (HD), a common distance metric, calculates the maximum distance between boundary points of segmentation results and the gold - standard to evaluate boundary fitting. A smaller value indicates more accurate boundary segmentation and reflects spatial positioning accuracy.

The Average Surface Distance (ASD) calculates the average distance between boundary points, better reflects the overall boundary fitting, and avoids the impact of extreme values.

The Volume Similarity (VS) evaluates volume consistency between segmentation results and the gold - standard by calculating the ratio of the absolute volume difference to the gold - standard volume, and suits clinical scenarios with high volume - measurement accuracy requirements.

These metrics quantify segmentation quality from different dimensions and are usually used together in research to comprehensively evaluate model performance. Some clinical studies also involve subjective evaluation by medical experts, combining technical indicators and clinical practicality to assess the application value of segmentation technology.

5. Existing Problems and Challenges

5.1 Scarcity of Annotated Data

The scarcity of annotated data is the core bottleneck restricting the application of machine vision in medical image segmentation. Medical image annotation requires a high degree of professionalism and must be completed by experienced radiologists, while qualified medical expert resources are limited. In addition, the annotation process is time-consuming and labor-intensive. The detailed annotation of a 3D medical image often takes several hours or even days, making it difficult to carry out on a large scale.

Privacy protection and data sharing restrictions further exacerbate the problem of data scarcity. Medical images contain patients' sensitive information, and are restricted by medical privacy regulations. Data from different hospitals and regions cannot be freely shared, resulting in scattered data distribution and difficulty in forming large-scale annotated datasets. Even within the same hospital, annotation standards may vary among different departments and physicians, leading to difficulty in ensuring the consistency of data annotation and affecting model training effects.

Data distribution imbalance also brings challenges to model training. In clinical practice, there are relatively more imaging data of common diseases, while imaging data of rare diseases is extremely scarce; data of large lesions is relatively abundant, while data of small lesions is scarce. This unbalanced distribution leads the model to tend to learn the features of common cases during training, resulting in low segmentation accuracy for rare cases or small lesions, which is difficult to meet the diverse clinical needs.

5.2 Inconsistent Image Formats

Inconsistent medical image formats are a major obstacle to the clinical implementation of machine vision segmentation technology. Currently, clinical medical imaging equipment comes from different manufacturers (e.g., Siemens, General Electric, Philips), with differences in imaging parameters and scanning protocols, resulting in varying output image formats and quality. Although the DICOM format is widely used as the standard for medical images, DICOM files from different hospitals and equipment vary in tag information, pixel spacing, slice thickness and other parameters, making standardized image data processing difficult. For instance, CT value calibration differs among equipment, and the grayscale features of the same organ vary under different equipment, affecting the generalization ability of segmentation models.

Besides format differences, individual image quality differences cannot be ignored. Factors like patients' body shape, scanning position and respiratory movement can cause artifacts and uneven grayscale in images. Variances in equipment maintenance levels and scanning technicians' operating standards in different hospitals further worsen image quality fluctuations. These factors lead to a significant drop in segmentation accuracy when a segmentation model trained on data from one hospital is applied to data from other hospitals, making cross - center and cross - equipment promotion and application difficult.

5.3 Other Challenges

In addition to the above core problems, machine vision segmentation technology also faces challenges such as insufficient model generalization ability and poor clinical interpretability. Most existing segmentation models are trained on public datasets or data from a single hospital, which are highly sensitive to changes in data distribution and difficult to adapt to complex and variable image scenarios in clinical practice, such as rare lesions and anatomical structure variations.

Poor clinical interpretability is an inherent flaw of deep learning models. Most existing segmentation models are "black box" models, which cannot clearly explain the basis for segmentation decisions, leading to insufficient trust of physicians in segmentation results and difficulty in fully relying on models for clinical decisions. In addition, the adaptability of segmentation technology to clinical workflows needs to be improved. Most existing models are independent analysis tools, which are difficult to seamlessly integrate with hospital PACS systems and electronic medical record systems, affecting clinical application efficiency.

6. Progress in Clinical Integration

6.1 Diagnostic Auxiliary Scenarios

Machine vision segmentation technology has achieved initial integration in clinical diagnosis, becoming an important auxiliary tool for physicians. In the diagnostic process of the radiology department, segmentation models can automatically extract organ and lesion regions, generate quantitative analysis reports, help physicians quickly locate lesion positions, evaluate lesion ranges, and reduce the risk of missed diagnosis and misdiagnosis.

For example, in brain MRI image diagnosis, segmentation models can automatically identify brain tumor regions, calculate tumor volume and edema range, and label the positional relationship between tumors and surrounding important brain tissue and blood vessels, providing intuitive diagnostic references for physicians. For physicians in primary hospitals or those with insufficient experience, such auxiliary tools can significantly improve diagnostic accuracy and narrow the diagnostic level gap between different medical institutions.

In the diagnosis of abdominal diseases, automatic segmentation and quantitative analysis of the liver and kidneys can assist physicians in evaluating organ function status. For example, parameters such as liver volume and surface roughness of patients with cirrhosis can be automatically calculated through segmentation technology, providing objective basis for liver function classification; segmentation results of the kidneys in patients with kidney stones can help physicians accurately judge the position and size of stones and formulate treatment plans.

6.2 Treatment Planning and Efficacy Evaluation

The integrated application of segmentation technology in treatment planning and efficacy evaluation is more in-depth. In tumor radiotherapy, segmentation models can accurately delineate tumor target volumes and surrounding normal organs, providing precise anatomical structure data for radiotherapy planning systems, ensuring that radiotherapy rays accurately cover the tumor region, and maximizing the protection of normal tissues, thereby reducing radiotherapy side effects.

In surgical treatment, segmentation technology provides important support for preoperative planning. For example, before surgery for liver tumor patients, by segmenting the liver, tumor, and blood vessel structures from CT images, physicians can simulate the surgical resection process on a 3D reconstructed model, optimize the resection range, evaluate the volume and function of the remaining liver, and improve surgical success rate. In kidney transplantation surgery, automatic kidney segmentation and vascular reconstruction can help physicians evaluate the quality of donor kidneys, plan surgical anastomosis paths, and shorten surgical time.

In efficacy evaluation, segmentation technology can realize quantitative monitoring of treatment effects. By comparing lesion segmentation results before and after treatment, parameters such as tumor volume change rate and lesion density change can be accurately calculated, providing objective indicators for efficacy evaluation. For example, after interventional therapy for liver cancer, regular segmentation of tumor regions can quickly judge whether the tumor is shrinking or recurring, and adjust the treatment plan in a timely manner.

6.3 Multi-Center Clinical Verification

Multi-center clinical verification is a key link for segmentation technology to move towards large-scale application. In recent years, many hospitals and research institutions have collaborated to carry out multi-center clinical trials of segmentation technology, verifying the applicability and stability of models in different medical scenarios.

In multi-center verification, segmentation models need to process image data from different hospitals and equipment, and evaluate model generalization ability through unified standardized processes and evaluation indicators. For example, the liver tumor segmentation model trained based on the LiTS dataset has shown good cross-center adaptability in clinical verification in many hospitals across the country, with the mean DSC of segmentation accuracy and medical expert annotation results reaching above 0.85, meeting clinical application needs.

Multi-center clinical verification not only verifies the clinical value of segmentation technology, but also provides important data support for technical optimization. By analyzing the verification results of different centers, researchers can identify the weak links of the model, such as insufficient segmentation accuracy of small lesions and poor processing effect of images from specific equipment, and targeted improve model algorithms to enhance the clinical adaptability of technology.

7. Conclusion and Outlook

Machine vision segmentation technology has great application potential in medical image processing and is an important clinical diagnosis and treatment auxiliary. This paper reviews organ and lesion segmentation of CT and MRI images, analyzing technical basis, application scenarios, datasets, evaluation indicators, existing problems, and clinical integration progress. Research shows it can improve medical image analysis efficiency and accuracy, offering objective references for organ function evaluation, etc., and has initial clinical applications in diseases like brain and liver tumors.

Currently, the technology faces challenges such as scarce annotated data, inconsistent image formats, and insufficient model generalization, restricting large - scale clinical implementation. In the future, breakthroughs are needed in data, technology, and clinical practice. At the data level, use weakly and semi - supervised learning to reduce data dependence and establish cross - center data sharing platforms. At the technical level, develop multi - modal fusion segmentation models[9]. At the clinical level, strengthen integration with clinical workflows and develop embedded tools for seamless docking with hospital systems.

With technological progress and clinical needs, the technology will play a greater role in medicine, empowering the whole process from auxiliary diagnosis to precise treatment, supporting medical service quality improvement and cost reduction. Future research should be more clinically oriented, strengthen medical - engineering cooperation, and promote the technology from the laboratory to clinical practice to serve patients and physicians.

References

[1] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Cham: Springer international publishing, 2015.

[2] Menze, Bjoern H., et al. "The multimodal brain tumor image segmentation benchmark (BRATS)." IEEE transactions on medical imaging 34.10 (2014): 1993-2024.

[3] Bilic, Patrick, et al. "The liver tumor segmentation benchmark (lits)." Medical image analysis 84 (2023): 102680.

[4] Milletari, Fausto, Nassir Navab, and Seyed-Ahmad Ahmadi. "V-net: Fully convolutional neural networks for volumetric medical image segmentation." 2016 fourth international conference on 3D vision (3DV). IEEE, 2016.

[5] Bakas, Spyridon, et al. "Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features." Scientific data 4.1 (2017): 1-13.

[6] Zhao, Hengshuang, et al. "Pyramid scene parsing network." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[7] Chen, Liang-Chieh, et al. "Rethinking atrous convolution for semantic image segmentation." arxiv preprint arxiv:1706.05587 (2017).

[8] Zhou, Huajun, et al. "Multimodal data integration for precision oncology: Challenges and future directions." arxiv preprint arxiv:2406.19611 (2024).

[9] Wang, Xizhao, Yanxia Zhao, and Farhad Pourpanah. "Recent advances in deep learning." International Journal of Machine Learning and Cybernetics 11.4 (2020): 747-750.