ACCSCIENCE PUBLISHING

Artificial Intelligence in Health

ORIGINAL RESEARCH ARTICLE

Enhancing spinal MRI segmentation with an asymmetric U-Net architecture

Longfei Zhou[1†*], Xingyu Chen[2†], Weihao Cheng[3], Zhanghao Qin[2], Tianao Shen[4], Pingyu Cao[5], Zebo Huang[2], Xiangyu Wu[6], and Yiyao Zhang[7]

[1]Department of Biomedical, Industrial and Systems Engineering, College of Engineering and Business, Gannon University, Erie, Philadelphia, United States of America

[2]School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi, Jiangsu, China

[3]School of Physics and Optoelectronic Engineering, Nanjing University of Information Science and Technology, Nanjing, Jiangsu, China

[4]School of Internet of Things Engineering, Jiangnan University, Wuxi, Jiangsu, China

[5]School of Remote Sensing and Surveying Engineering, Nanjing University of Information Science and Technology, Nanjing, Jiangsu, China

[6]School of Mechanical Engineering, Jiangnan University, Wuxi, Jiangsu, China

[7]College of Robot and Engineering, Guangzhou City University of Technology, Yinchuan, Ningxia, China

[†]*These authors contributed equally to this work.*

***Corresponding author:**
Longfei Zhou
(zhou009@gannon.edu)

Citation: Zhou L, Chen X, Cheng W, *et al*. Enhancing spinal MRI segmentation with an asymmetric U-Net architecture. *Artif Intell Health*. 2025;2(1):42-52. doi: 10.36922/aih.3889

Received: June 7, 2024

1st revised: July 5, 2024

2nd revised: July 12, 2024

Accepted: August 1, 2024

Published Online: October 21, 2024

Copyright: © 2024 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Abstract

Spinal diseases are among the most prevalent health issues in modern society, significantly impacting patients' quality of life. Diagnosing conditions such as disc herniation and spinal deformity requires advanced medical imaging techniques, including X-rays, magnetic resonance imaging (MRI), computed tomography, and nuclear magnetic resonance. Spine MRI is particularly crucial due to its ability to provide high-resolution images of soft tissues, essential for accurate diagnosis. However, the manual segmentation of spine MRI images is labor-intensive and inadequate for large-scale quantitative analysis. Thus, developing automated spinal MRI segmentation methods is critical to alleviating doctors' workload and enhancing diagnostic efficiency. In this study, we propose a novel asymmetric U-Net architecture designed to improve the precision of reconstructing complex structures and details by increasing the depth of the upsampling side. The model incorporates adjacent-scale skip connections to control parameters while maintaining high segmentation accuracy. In addition, residual connections on the upsampling side prevent gradient vanishing, thereby enhancing the network's feature learning and representation capabilities. Experimental results indicate that this method significantly reduces training time and increases model accuracy compared to traditional approaches, marking a substantial advancement in automated spinal MRI segmentation. This innovative approach holds promise for improving clinical outcomes and optimizing the workflow in medical imaging departments.

Keywords: Spinal magnetic resonance imaging; Automated segmentation; Asymmetric U-Net; Medical imaging; Deep learning

1. Introduction

Spinal diseases are among the most prevalent health issues in modern society. The pathogenesis of spine diagnostics involves understanding the underlying mechanisms and origins of spinal disorders, which is crucial for accurate diagnosis and effective treatment. Modern spine diagnostics has evolved significantly with advancements in imaging technology, genetic research, and molecular biology, providing deeper insights into spinal pathologies. Pathogenesis in spine diagnostics refers to the study of how spinal diseases develop and progress. This includes degenerative diseases such as osteoarthritis and intervertebral disc degeneration, as well as inflammatory conditions such as ankylosing spondylitis.[1] Degenerative spinal disorders often involve the breakdown of intervertebral discs and facet joints. Factors such as aging, mechanical stress, and genetic predisposition contribute to the degeneration process. Studies have shown that mechanical loading and biochemical changes play significant roles in disc degeneration.[2] Diagnostic tools such as magnetic resonance imaging (MRI) and computed tomography (CT) scans allow for detailed visualization of these changes.[3] Inflammatory spinal diseases, such as ankylosing spondylitis, involve chronic inflammation of the spinal joints, which leads to pain and stiffness. The pathogenesis of these diseases is linked to genetic markers such as *HLA-B27*.[4] Advanced diagnostic methods, including MRI and blood tests for inflammatory markers, are essential for early detection and monitoring.[5] Spinal tumors' pathogenesis includes genetic mutations and environmental factors that lead to abnormal cell growth. Diagnostic imaging, biopsy, and molecular testing are crucial in identifying and characterizing spinal tumors, guiding treatment decisions.[6,7]

The advancements in spine diagnostics have a profound impact on modern healthcare, influencing clinical practice, patient outcomes, and healthcare systems. Improved imaging technologies, such as high-resolution MRI and three-dimensional (3D) CT scans, provide detailed visualization of spinal structures, enhancing diagnostic accuracy.[8,9] The integration of genetic and molecular diagnostics enables personalized treatment plans. By understanding the genetic and molecular basis of spinal diseases, clinicians can tailor therapies to individual patients, improving efficacy and reducing adverse effects.[10] Advances in diagnostic imaging have facilitated the development of minimally invasive surgical techniques. Real-time imaging guidance during procedures minimizes tissue damage, reduces recovery time, and lowers the risk of complications.[11] Early detection of spinal disorders through advanced diagnostics allows for timely intervention, potentially preventing disease progression and reducing the burden of chronic spinal conditions on patients and healthcare systems. Studies have shown that early intervention can significantly improve long-term outcomes for patients with spinal conditions.[12] Accurate diagnostics and early interventions can reduce healthcare costs by decreasing the need for extensive surgeries and long-term care. Efficient diagnostic processes also streamline patient management, optimizing resource utilization within healthcare systems.

Accurate diagnosis of spinal conditions is critically dependent on the analysis provided by MRI.[9,13-14] High-quality spine MRI segmentation is crucial for enabling doctors to precisely locate and examine spinal structures, thereby facilitating the diagnosis of various spine-related diseases such as disc herniation and spinal deformities.[15,16] Accurate segmentation results are essential for assessing the severity of these conditions and developing effective treatment plans.

To address the labor-intensive nature of medical imaging tasks, there is a growing trend toward data-driven approaches in contemporary medical imaging technology.[17,18] Traditional image processing techniques, such as thresholding,[19] edge detection,[20] and mathematical morphology,[21] have yielded some positive outcomes. However, significant advancements have been made in recent years with the advent of deep learning methods, particularly convolutional neural networks, which have revolutionized spine medical image segmentation.[22,23] Models such as U-Net,[24] DeepLab,[25] and Fully Convolutional Networks[26] have been extensively applied to spine image segmentation tasks, with the U-Net architecture achieving notable success. This model excels in extracting and representing feature information in MRI medical image analysis.

To further enhance U-Net's ability to capture multi-scale information, Huang *et al.* introduced U-Net++,[27] which incorporates full-scale skip connections. This design effectively aggregates low-level and high-level semantic information, improving segmentation performance. However, the practical application of U-Net++ is challenged by the large size of spine medical images and the network's complex structure, leading to prolonged training times and reduced accuracy due to the intricate nature of spinal images.

This study proposes an innovative and efficient spine segmentation method called "J-Unet" network to overcome these limitations and improve model accuracy while reducing training costs. Our approach includes optimizing multi-scale skip connection paths and deepening the network depth of the upsampling component to capture

finer features across different depths. The key contributions of this study are as follows:

(1) Optimization of multi-scale skip connection paths: By refining these paths, we can better control model parameters while ensuring high segmentation accuracy.
(2) Increased network depth in the upsampling process: This modification enables the model to reconstruct complex structures and details more accurately, enhancing overall precision.
(3) Introduction of residual connections: These connections are incorporated locally to mitigate the problem of gradient vanishing, accelerate training, and strengthen the network's feature learning and representation capabilities.

In summary, the proposed method addresses the challenges of excessive training time and model accuracy in spine MRI segmentation. By optimizing the architecture of U-Net++, we aim to provide a more efficient and precise tool for medical professionals, thereby improving diagnostic and treatment outcomes for spinal diseases.

2. Data and methods

In this section, the dataset used in this study and the processing methods employed are introduced. A detailed description of the proposed J-Unet architecture, designed specifically for spine MRI image segmentation, is also provided.

2.1. Dataset and processing methods

The dataset utilized in this study consists of T2-weighted MRI scans obtained from a cohort of 215 patients, sourced from multiple medical institutions to ensure diversity and robustness in the data.[28,29] All images are provided in Nifti format, a widely recognized standard for medical imaging, which facilitates comprehensive 3D visualization and analysis.

Initially, the dataset's labels included 21 distinct pixel values representing various anatomical structures and features. For the purposes of this study, these original labels were reclassified into three primary categories: vertebral bone, intervertebral disc, and background regions. This reclassification simplifies the segmentation task, focusing on the most clinically relevant structures. The reclassified pixel values for segmentation are detailed in Table 1, providing a clear framework for subsequent image processing and analysis.

We configured the training and validation datasets with a 4:1 ratio. To ensure a balanced and representative distribution of data, the allocation was performed through random sampling, which mitigates potential biases and ensures that the training dataset captures a comprehensive range of variations present in the images. Each input spine image is standardized to a size of 512 × 512 pixels to maintain consistency and optimize the computational efficiency during model training. Figure 1 shows an example of an original spinal MRI image and its corresponding label image, illustrating the visual clarity and distinct boundaries of the segmented regions.

Table 1. Reclassified pixel values for segmentation

Vertebral bone	Intervertebral disc	Background regions
(100 100 100)	(255 255 255)	(0 0 0)

2.2. J-Unet architecture

The J-Unet architecture introduced in this study represents a significant evolution from traditional U-Net and U-Net3+ designs, addressing some of the limitations inherent in these models and incorporating advanced features to enhance performance in spine MRI image segmentation. Figure 2 illustrates the architecture of the proposed J-Unet, highlighting the innovative structural elements that differentiate it from its predecessors.

2.2.1. Asymmetric network architecture

One of the most notable advancements in the J-Unet model is its asymmetric network architecture, which diverges from traditional U-Net structures by extending the upsampling pathway with three additional layers compared to its downsampling counterpart. This asymmetry is deliberately engineered to enhance the network's ability to capture and reconstruct complex, hierarchical features more comprehensively. By increasing the depth on the upsampling side, the model can progressively refine the spatial resolution of the feature maps, thereby improving the accuracy of the segmentation output.

In this design, the input and output channels of these additional upsampling layers are set to 120, effectively reducing the number of parameters without compromising accuracy. The advantages of an asymmetric network architecture are evident in its enhanced ability to precisely reconstruct complex structures and details. Traditional symmetric architectures often struggle with images characterized by irregular shapes and intricate edges. By increasing the depth on the upsampling side, the asymmetric network better adapts to these complexities. It also utilizes skip connections between specific scales to fine-tune model parameters while maintaining segmentation accuracy.

The depth of the upsampling pathway in the J-Unet is crucial for reconstructing the finer details of spinal

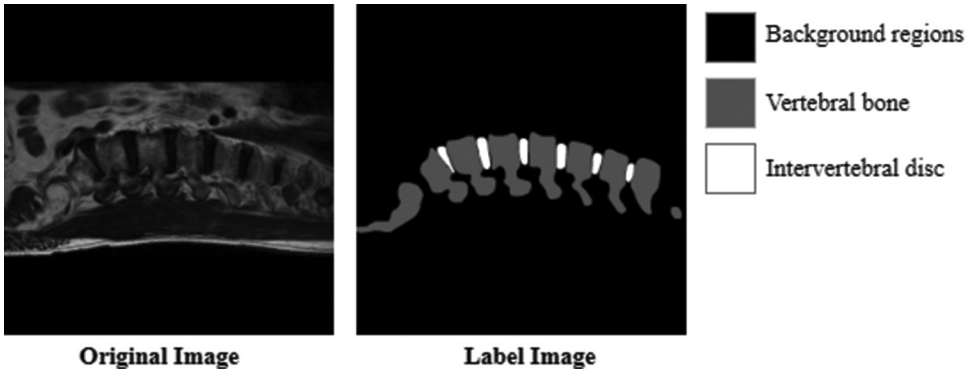

Figure 1. Example original spinal magnetic resonance imaging image and its corresponding label image

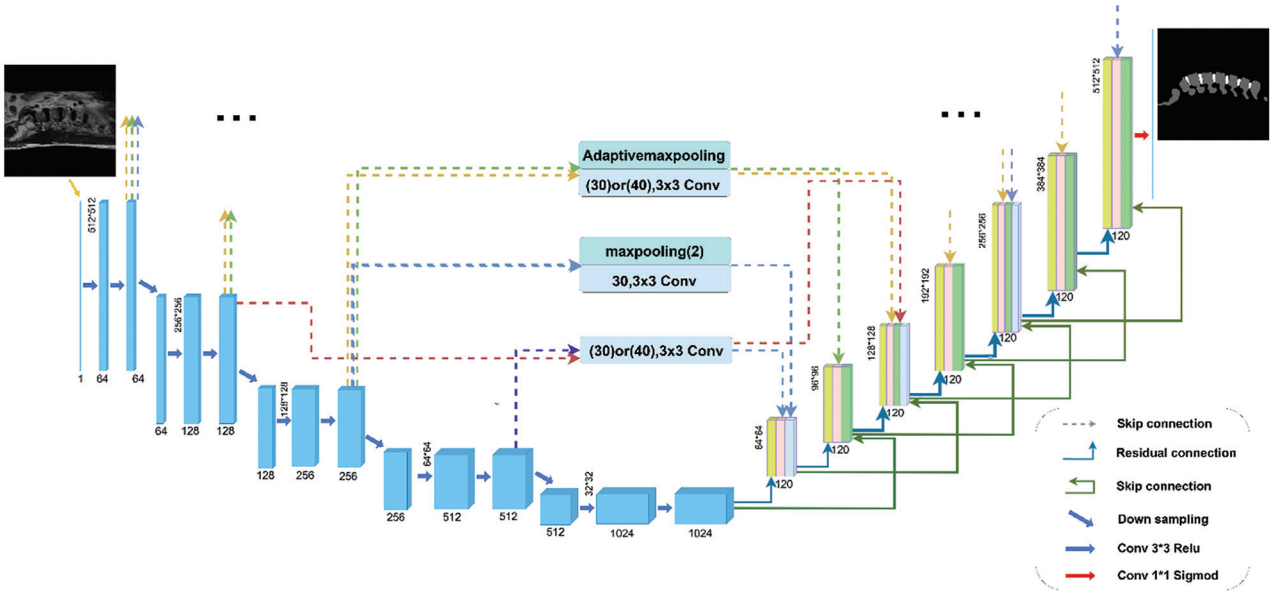

Figure 2. The architecture of the proposed J-Unet

structures, which are often lost in traditional symmetric architectures. This extended pathway allows the network to learn more detailed and nuanced representations of the spinal structures, crucial for tasks that require high precision, such as medical image segmentation. The ability to capture and reconstruct hierarchical features ensures that the segmentation of the spine is both accurate and reliable, which is essential for clinical applications where precise anatomical delineation is required for diagnosis and treatment planning.

2.2.2. Adjacent-scale skip connections

Unlike the U-Net3+ architecture, which utilizes fully-scaled skip connections, the J-Unet adopts adjacent-scale skip connections. This approach strategically reduces the redundancy and overall parameter count associated with the skip connections while maintaining the ability to leverage multi-scale feature information effectively. Fully-scaled skip connections can introduce excessive redundancy, leading to an unnecessary increase in computational load and model complexity. In contrast, adjacent-scale skip connections streamline the network, reducing computational overhead without sacrificing the richness of the multi-scale features. This optimization not only simplifies the network architecture but also enhances computational efficiency and model scalability, making the J-Unet more practical for large-scale applications and real-time processing.

In the J-Unet model, the size of feature maps does not uniformly change by integer multiples. For instance, adaptive max pooling is employed to resize a 256 × 256 feature map from the encoder to a 192 × 192 feature map in the decoder, whereas standard fixed window max pooling is used for regular size adjustments. This method enhances

internal information flow and integration within the network, improves the perception of detailed structures and edges, and ultimately boosts the accuracy of segmentation tasks. Figure 3 illustrates these tailored connections in the J-Unet architecture, showcasing how the model efficiently integrates multi-scale information without the overhead of fully-scaled connections.

Adjacent-scale skip connections allow the network to capture both fine-grained details and coarse-grained semantics across different scales but with fewer parameters. This design retains the ability to capture detailed features necessary for accurate segmentation while reducing the computational burden, making the model more efficient and scalable. The strategic use of these connections ensures that the model can effectively integrate information from different scales, enhancing its ability to accurately segment complex spinal structures.

2.2.3. Partial residual connections (PRCs)

The addition of three upsampling layers not only deepens the network architecture but also introduces the risk of gradient vanishing. To counteract this and boost the network's feature representation capabilities, residual connections are strategically employed during the upsampling process. Residual connections, introduced by He *et al.*[30] in their seminal work on deep residual networks, are designed to preserve and reuse the information captured in earlier layers, addressing the vanishing gradient problem and facilitating the training of deeper networks.

In the J-Unet model, each decoder layer is connected to the largest-scale feature map from the most adjacent layer through a residual connection, allowing for the construction of an even deeper network. The innovation of PRCs offers several advantages. Unlike traditional global residual connections, PRCs are more selective, maintaining and transmitting essential feature information and thereby enhancing the network's ability to learn complex feature representations more efficiently.

This selective connection strategy not only accelerates the training process but also reduces resource consumption. Overall, PRCs significantly improve training stability, enhance learning capacity, and bolster the generalization performance of neural networks, making them a vital component in the design of advanced deep learning models. By preserving critical information from earlier layers and reintroducing it at later stages, PRCs facilitate a more robust learning process, enabling the network to capture intricate details and complex patterns within the spinal MRI images.

2.2.4. Integration of advanced structural elements

The integration of these advanced structural elements in J-Unet, including extended asymmetry in the upsampling path, optimized skip connections, and strategic use of residual connections, aims to improve the accuracy and efficiency of spine MRI image segmentation. These enhancements enable the model to handle the intricacies of medical imaging data, which often involve complex anatomical variations and subtle pathological features.

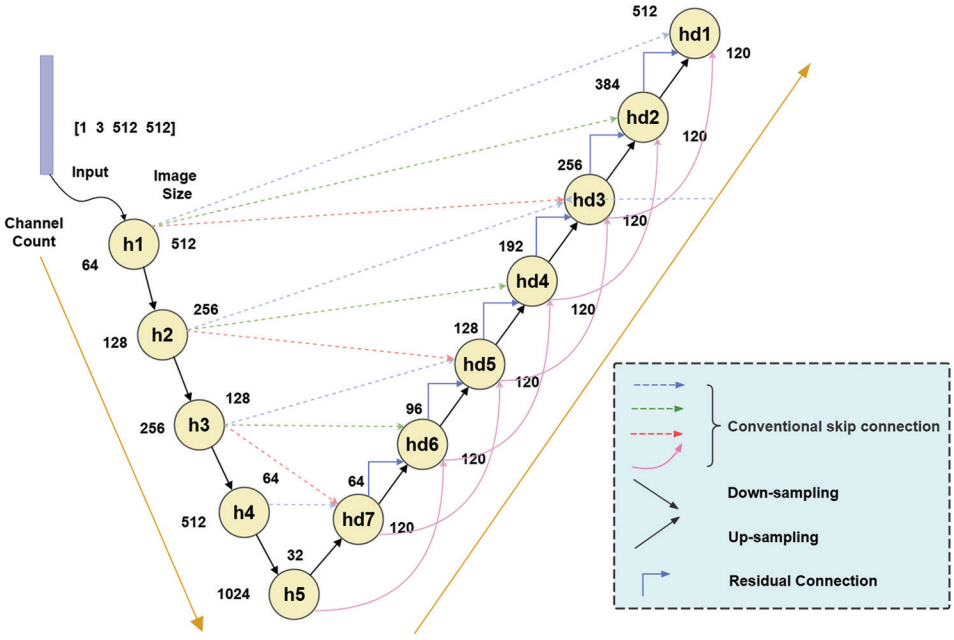

Figure 3. Detailed illustration of connections in J-Unet

By effectively managing these complexities, the J-Unet architecture supports more precise diagnostic outcomes and potentially informs better clinical decision-making.

Accurate segmentation of spine MRI images can aid in the early detection and diagnosis of spinal pathologies such as disc herniation, spinal stenosis, and tumors. Precise delineation of these structures is essential for planning surgical interventions, assessing disease progression, and evaluating treatment efficacy. The improved segmentation performance offered by the J-Unet model can therefore contribute to better patient outcomes by enabling more targeted and effective treatments.

Moreover, the computational efficiency and scalability of the J-Unet architecture make it suitable for deployment in clinical settings where rapid processing of large volumes of imaging data is required. This is particularly important in modern healthcare environments, where the demand for advanced imaging techniques is increasing, and the ability to process and analyze data quickly can significantly impact the quality of care provided.

3. Results and analysis

In this section, we detail the organization and methodology of our experimental study using magnetic resonance (MR) imaging data. The dataset, comprising MR image sequences from 215 patients, was stratified into training and test sets in a 4:1 ratio. This separation was carefully designed to ensure both sets were representative of the overall dataset, supporting the generalizability of our findings. Using the J-Unet architecture, we developed a model capable of automatically segmenting vertebrae and intervertebral discs within spinal MR images. The performance of this model was rigorously evaluated to ascertain its efficacy in medical imaging tasks. In addition, to provide a comprehensive analysis of our model's capabilities, we compared its performance with several other established neural network architectures: Unet, Unet++, Unet+++, and Res-Unet. This comparative study aimed to highlight the strengths and potential areas for improvement in the J-Unet architecture relative to other models in handling complex segmentation tasks in spinal MR images.

The computational experiments were conducted using the Pytorch framework and cuDNN library, optimized for deep neural network operations. All models were trained on a robust hardware setup featuring a single NVIDIA GTX 3090 graphics processing unit. This high-performance computing environment ensured efficient processing of large-scale data, facilitating timely training and evaluation of the models. This systematic approach allowed us to not only assess the specific advantages of the J-Unet model but also to establish a baseline for performance against other prominent architectures in the field, thereby providing a clear perspective on the current state of the art in spinal MR image segmentation.

3.1. Experimental setup

We utilized three evaluation metrics to assess the segmentation performance of our models: accuracy, mean intersection over union (mIOU), and dice coefficient. These metrics provide a comprehensive view of the model's effectiveness in segmenting vertebrae and intervertebral discs in spinal MR images. The formulas for these metrics are as follows:

(1) Accuracy: Measures the proportion of true results (both true positives [TP] and true negatives [TN]) among the total number of cases examined.

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \tag{I}$$

(2) mIOU: Calculates the average IOU across all classes, providing an overall measure of segmentation performance.

$$mIOU = \frac{cIOU + ucIOU}{2} \tag{II}$$

(3) Dice coefficient: Evaluates the overlap between the predicted and ground truth segments, often used in medical image analysis.

$$Dice = \frac{2TP}{2TP + FP + FN} \tag{III}$$

Definitions of confusion matrix terms are given below:

- TP: A positive class instance correctly predicted as positive.
- False negative (FN): A positive class instance incorrectly predicted as negative.
- False positive (FP): A negative class instance incorrectly predicted as positive.
- TN: A negative class instance correctly predicted as negative.

In these definitions, TP and TN indicate correct predictions of the instance class, while FN and FP indicate incorrect predictions. Table 2 outlines the specific parameter settings used for the segmentation model.

These settings were carefully chosen to optimize model performance and ensure reliable evaluation metrics. They

Table 2. Parameter settings of the segmentation model

Parameter	Learning rate	Optimizer	Batch size	Epoch
Value	1e-4	Adam	2	100

include details on learning rate, batch size, number of epochs, and other hyperparameters critical for training deep learning models. The segmentation performance of J-Unet was benchmarked against several established models: Unet, Unet++, Unet+++, and Res-Unet. Each model was trained and evaluated using the same dataset and parameter settings to ensure a fair comparison. The results were analyzed using the aforementioned metrics to determine the relative strengths and weaknesses of each model.

3.2. Optimization and loss function selection

In this study, we utilized the Adam optimizer due to its significant advantages in deep learning model training. Adam's ability to apply different scaling factors to parameter updates facilitates the discovery of the global optimum more efficiently during the training process. This tailored approach to parameter adjustment helps navigate the complex loss landscapes often encountered in deep learning models, ensuring that each parameter evolves at an appropriate pace. By adaptively adjusting the learning rate for each parameter, Adam accelerates convergence, enhancing the overall efficiency of the training process. In addition, Adam is known for its computational efficiency and ease of implementation. It requires less memory compared to other optimizers, making it ideal for handling large-scale data and complex models. To ensure stable convergence towards a local optimum, we set the learning rate to 0.0001. This small learning rate helps in fine-tuning the model parameters, preventing overshooting and ensuring a smooth descent in the loss landscape.

The choice of Dice Loss as the loss function for this study is driven by its effectiveness in addressing common challenges in medical image segmentation. Dice Loss is particularly robust in scenarios with imbalanced classes, a common occurrence in medical imaging. Its calculation involves the intersection and union of the predicted and true values, making it less sensitive to the disproportionate pixel counts of different classes. This robustness ensures that the model performs well even when certain classes are underrepresented.

Moreover, Dice Loss provides smoother gradients compared to other loss functions, contributing to a more stable training process. This stability helps mitigate issues such as exploding or vanishing gradients, which can hinder the training of deep neural networks. By emphasizing the similarity between predicted and true segmentations, Dice Loss encourages the model to produce accurate and refined segmentation results. This focus on overlap and accuracy is crucial for tasks requiring precise delineation of structures, such as in medical imaging where detail and precision are paramount.

The proven applicability of Dice Loss in medical imaging further validates its use in this study. Dice Loss is widely employed in medical image segmentation tasks, particularly those demanding high accuracy, such as tumor segmentation. Its sensitivity to fine structures and boundaries makes it an excellent choice for medical applications, where the accurate segmentation of anatomical structures is critical for diagnosis and treatment planning. The combination of the Adam's optimizer and Dice Loss provides a robust framework for training our segmentation model. Adam's efficiency and adaptability, coupled with Dice Loss's robustness to class imbalance and emphasis on accurate segmentation, ensure a high-performance model suitable for the complexities of medical image analysis.

3.3. Model training

During the training process, Dice Loss was employed as the loss function for our deep learning model. Dice Loss is a widely used metric in image segmentation tasks, designed to measure the similarity between the predicted segmentation and the ground truth.

The segmentation models, namely UNET, UNET++, UNET+++, ResUNET, and J-UNET, were trained using the training set. Figure 4 illustrates the variation of loss and Dice coefficient during training and validation for these five models. The graphs provide a clear comparison of how each model's performance evolved over the training epochs. During the training process, we encountered several instances of gradient explosions, particularly with the UNET+++ model. These gradient explosions resulted in a sharp increase in loss values. To address this issue, we employed a strategy of resuming training from checkpoints saved before the occurrence of the gradient explosions. This approach allowed us to continue training effectively, leading to eventual convergence. The training of the remaining network models exhibited normal convergence patterns without such disruptions.

Upon analyzing the final convergence ranges, it was evident that J-UNET and UNET achieved lower loss values on the training set compared to the other models. On the test set, although the differences in loss values among the various models were not substantial, J-UNET and ResUNET consistently demonstrated smaller loss values. This indicates that J-UNET, in particular, exhibited superior generalization ability, aligning well with the task objectives. The combination of Dice Loss and the robust architecture of J-UNET contributed to its superior performance in segmenting vertebrae and intervertebral discs in spinal MR images. The J-UNET model not only converged effectively but also showed a strong ability to generalize from the training set to the test set, making it a highly effective tool for medical image segmentation tasks.

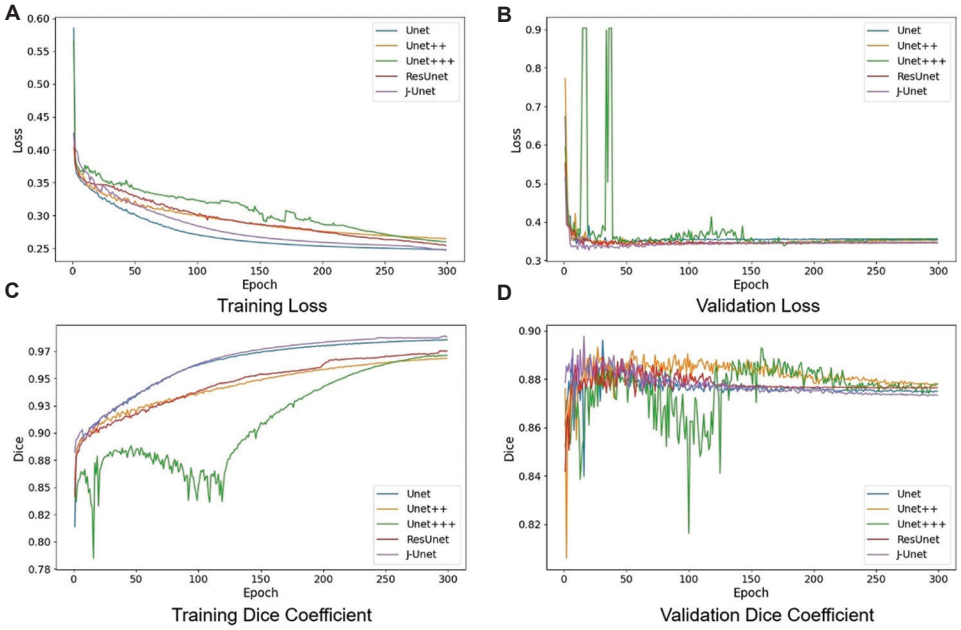

Figure 4. Loss values and evaluation metrics in model training and validation. (A) Training Loss; (B) Validation Loss; (C) Training Dice Coefficient; and (D) Validation Dice Coefficient

3.4. Results

Upon completion of the training phase for the various network models, we computed several key metrics on the test set to assess their performance. These metrics include dice, accuracy, and mIOU, which are standard measures for evaluating segmentation quality. For detailed formulas of these metrics, refer to section 3. The results, presented in Table 3, highlight the superior performance of the J-Unet model across all evaluation criteria.

As illustrated in Table 3, the J-Unet model achieved notable improvements in comparison to other models. Specifically, J-Unet improved the dice score by at least 0.24%, accuracy by 0.74%, and mIOU by at least 0.24%. These improvements underscore the effectiveness of the J-Unet architecture in accurately segmenting spinal MR images.

In addition to evaluating segmentation performance, we compared the number of parameters across the different models, as shown in Table 4. Our J-Unet model, while having a parameter count slightly higher than UNET, boasts approximately 78.1% fewer parameters than Res-UNET, 17.2% fewer than UNET+++, and roughly 2.3% fewer than Res-UNET. Despite its relatively smaller parameter count, J-Unet consistently achieved the highest performance in all evaluation metrics. This efficiency indicates that our model is not only accurate but also resource-effective.

The experimental results clearly demonstrate that the design choices of J-Unet, such as the asymmetric network structure, adjacent-scale skip connections, and PRCs, contribute significantly to its enhanced accuracy and generalization capabilities. These architectural innovations enable J-Unet to maintain high performance with fewer parameters, reducing the computational resources and time required for both training and inference. Figure 5 shows the segmentation maps of different models. The superior performance and efficiency of J-Unet affirm the benefits of its innovative design. The model's ability to achieve high segmentation accuracy with a smaller parameter footprint makes it valuable for medical image

Table 3. Scores obtained of all models

Model	Dice	Accuracy	mIOU
Unet	0.8791	0.9613	0.8093
Unet++	0.8885	0.9655	0.8264
Unet+++	0.8866	0.9606	0.8248
Res-Unet	0.8889	0.9642	0.8241
J-Unet	**0.8913**	**0.9729**	**0.8288**

The values in boldface indicate the best performance among all models in each metric.
Abbreviation: mIOU: Mean intersection over union.

Table 4. The number of parameters of different models

Model	Unet	Unet++	Unet+++	Res-Unet	J-Unet
Total parameters	17267523	24423232	25659999	101942977	22331979

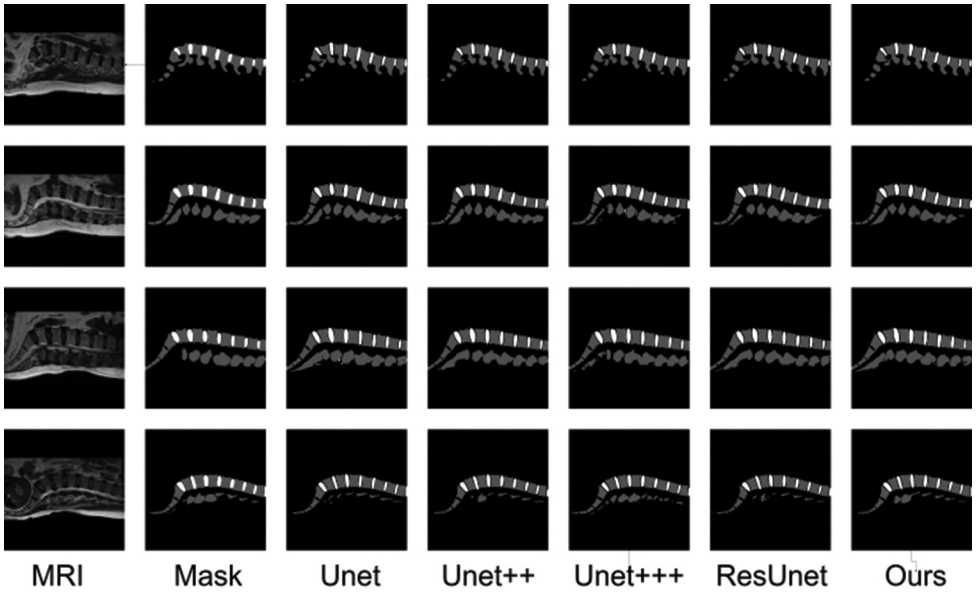

Figure 5. Segmentation maps of different models

segmentation, providing a practical solution that balances performance with computational efficiency.

4. Conclusion

In this study, we introduce a novel asymmetric U-Net architecture, termed J-Unet, designed for efficient and accurate spine MRI image segmentation. J-Unet distinguishes itself from conventional U-Net and its variants through its asymmetric encoder-decoder structure, featuring a deeper upsampling path that enhances the precise reconstruction of anatomical details. The incorporation of adjacent-scale skip connections and PRCs allows J-Unet to reduce the number of model parameters while maintaining the flow of multi-scale contextual information. We rigorously evaluated J-Unet using a dataset comprising 215 spine MRI images. The results indicate that J-Unet significantly outperforms other models, including U-Net, U-Net++, and U-Net+++, achieving at least a 0.24% improvement in both dice score and mIoU. Furthermore, J-Unet operates with substantially fewer parameters compared to U-Net+++ and Res-UNET, demonstrating superior performance and efficiency. In conclusion, this study presents an innovative asymmetric U-Net architecture specifically tailored for spine MRI image segmentation. J-Unet's unique design enables precise localization and segmentation while optimizing parameter efficiency, making it a highly accurate and resource-effective solution. Our model offers a promising advancement in automating spine segmentation in medical image analysis, potentially enhancing diagnostic processes and treatment planning.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Longfei Zhou

Formal analysis: Xingyu Chen, Weihao Cheng, Zhanghao Qin, Tianao Shen, Pingyu Cao, Zebo Huang, Xiangyu Wu, Yiyao Zhang

Investigation: Longfei Zhou, Xingyu Chen

Methodology: Xingyu Chen, Weihao Cheng, Zhanghao Qin, Tianao Shen, Pingyu Cao, Zebo Huang, Xiangyu Wu, Yiyao Zhang

Writing–original draft: Xingyu Chen, Weihao Cheng, Zhanghao Qin, Tianao Shen, Pingyu Cao, Zebo Huang, Xiangyu Wu, Yiyao Zhang

Writing–review & editing: Longfei Zhou

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

Data are available from the corresponding author upon reasonable request.

References

1. Kubaszewski Ł, Wojdasiewicz P, Rożek M, *et al*. Syndromes with chronic non-bacterial osteomyelitis in the spine. *Reumatologia*. 2015;53(6):328-336.

   doi: 10.5114/reum.2015.57639

2. Rajasekaran S, Bajaj N, Tubaki V, Kanna RM, Shetty AP. ISSLS prize winner: The anatomy of failure in lumbar disc herniation: An *in vivo*, multimodal, prospective study of 181 subjects. *Spine* (*Phila Pa 1976*). 2013;38(17):1491-1500.

   doi: 10.1055/s-0034-1376749

3. Adams MA, Roughley PJ. What is intervertebral disc degeneration, and what causes it? *Spine* (*Phila Pa 1976*). 2006;31(18):2151-2161.

   doi: 10.1097/01.brs.0000231761.73859.2c

4. Pedersen SJ, Maksymowych WP. The pathogenesis of ankylosing spondylitis: An update. *Curr Rheumatol Rep*. 2019;21(10):58.

   doi: 10.1007/s11926-019-0856-3

5. Aouad K, Maksymowych WP, Baraliakos X, Ziade N. Update of imaging in the diagnosis and management of axial spondyloarthritis. *Best Pract Res Clin Rheumatol*. 2020;34(6):101628.

   doi: 10.1016/j.berh.2020.101628

6. Hassan I, Wietfeldt ED. Presacral tumors: Diagnosis and management. *Clin Colon Rectal Surg*. 2009;22(2):84-93.

   doi: 10.1055/s-0029-1223839

7. Hashimoto K, Nishimura S, Miyamoto H, Toriumi K, Ikeda T, Akagi M. Comprehensive treatment outcomes of giant cell tumor of the spine: A retrospective study. *Medicine* (*Baltimore*). 2022;101(32):e29963.

   doi: 10.1097/MD.0000000000029963

8. Kalra MK, Maher MM, Toth TL, *et al*. Strategies for CT radiation dose optimization. *Radiology*. 2004;230(3):619-628.

   doi: 10.1148/radiol.2303021726

9. Winn A, Martin A, Castellon I, *et al*. Spine MRI: A review of commonly encountered emergent conditions. *Top Magn Reson Imaging*. 2020;29(6):291-320.

   doi: 10.1097/RMR.0000000000000261

10. Castaldo G, Lembo F, Tomaiuolo R. Molecular diagnostics: Between chips and customized medicine. *Clin Chem Labo Med*. 2010;48(7):973-982.

    doi: 10.1515/CCLM.2010.182

11. Arslantaş A, Dalbayrak S, Şimşek S, *et al*. Minimally Invasive Spine Surgery Current Aspects. Turkey: Ali Arslantaş. 2016.

12. Haldeman S, Kopansky-Giles D, Hurwitz EL, *et al*. Advancements in the management of spine disorders. *Best Pract Res Clin Rheumatol*. 2012;26(2):263-280.

    doi: 10.1016/j.berh.2012.03.006

13. Azimi P, Yazdanian T, Benzel EC, *et al*. A review on the use of artificial intelligence in spinal diseases. *Asian Spine J*. 2020;14(4):543.

    doi: 10.31616/asj.2020.0147

14. Da Costa RV, Moore SA. Differential diagnosis of spinal diseases. *Vet Clin North Am Small Anim Pract*. 2010;40(5):755-763.

    doi: 10.1007/978-981-16-9759-3_11

15. Cohen-Adad J, Alonso-Ortiz E, Abramovic M, *et al*. Generic acquisition protocol for quantitative MRI of the spinal cord. *Nat Protocols*. 2021;16(10):4611-4632.

16. Sollmann N, Löffler MT, Kronthaler S, *et al*. MRI-based quantitative osteoporosis imaging at the spine and femur. *J Magn Reson Imaging*. 2021;54(1):12-35.

    doi: 10.1002/jmri.27260

17. Willemink MJ, Koszek WA, Hardell C, *et al*. Preparing medical imaging data for machine learning. *Radiology*. 2020;295(1):4-15.

    doi: 10.1148/radiol.2020192224

18. Patel V. A framework for secure and decentralized sharing of medical imaging data via blockchain consensus. *Health Informatics J*. 2019;25(4):1398-1411.

    doi: 10.1177/1460458218769699

19. Senthilkumaran N, Vaithegi S. Image segmentation by using thresholding techniques for medical images. *Comput Sci Eng Int J*. 2016;6(1):1-13.

20. Song Y, Ma B, Gao W, Fan S. Medical image edge detection based on improved differential evolution algorithm and prewitt operator. *Acta Microscopica*. 2019;28(1).

21. Zhao F, Zhang J, Ma Y. Medical image processing based on mathematical morphology. In: *Proceedings of the 2012 International Conference on Computer Application and System Modeling (ICCASM 2012)*; 2012. p. 948-950.

    doi: 10.2991/iccasm.2012.241

22. Yamanakkanavar N, Choi JY, Lee B. MRI segmentation and classification of human brain using deep learning for diagnosis of Alzheimer's disease: A survey. *Sensors* (*Basel*). 2020;20(11):3243.

    doi: 10.3390/s20113243

23. Li H, Luo H, Huan W, *et al*. Automatic lumbar spinal MRI image segmentation with a multi-scale attention network. *Neural Comput Appl*. 2021;33:11589-11602.

    doi: 10.1007/s00521-021-05856-4

24. Ronneberger O, Fischer P, Brox T. *U-net: Convolutional Networks for Biomedical Image Segmentation*. Berlin: Springer; 2015. p. 234-241.

doi: 10.1007/978-3-319-24574-4_28

25. Hempe H, Yilmaz EB, Meyer C, Heinrich MP. Opportunistic CT screening for degenerative deformities and osteoporotic fractures with 3D DeepLab. In: *Medical Imaging 2022: Image Processing*. Bellingham, DC: SPIE; p. 127-134.

doi: 10.1117/12.2612848

26. Miao S, Piat S, Fischer P, *et al*. Dilated FCN for Multi-agent 2D/3D Medical Image Registration. In: *Proceedings of the AAAI Conference on Artificial Intelligence*.

doi: 10.1609/aaai.v32i1.11576

27. Huang H, Lin L, Tong R, *et al*. *Unet 3+: A full-Scale Connected UNet for Medical Image Segmentation*. United States: IEEE; p. 1055-1059.

doi: 10.1109/ICASSP40776.2020.9053405

28. Pang S, Pang C, Zhao L, *et al*. SpineParseNet: spine parsing for volumetric MR image by a two-stage segmentation framework with semantic image representation. *IEEE Trans Med Imaging*. 2020;40(1):262-273.

doi: 10.1109/TMI.2020.3025087

29. Pang S, Pang C, Su Z, *et al*. "DGMSNet: Spine segmentation for MR image by a detection-guided mixed-supervised segmentation network. *Med Image Anal*. 2022;75:102261.

doi: 10.1016/j.media.2021.102261

30. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2016. p. 770-778.

doi: 10.1109/CVPR.2016.90