Comparative Analysis of Deep Learning Models for Lumbar Vertebrae Segmentation in MRI Images

Kanstantsin Kurachka
Department of Information Technology
Pavel Sukhoi State Technical University of Gomel
Gomel, Belarus
kurochka@gstu.by

Huanhai Ren
Department of Information Technology
Pavel Sukhoi State Technical University of Gomel
Gomel, Belarus
rhh_cn@qq.com

Xuemei Wang
Department of Information Technology
Pavel Sukhoi State Technical University of Gomel
Gomel, Belarus
wxm_doc@qq.com

***Abstract.* Lower back pain is a widespread health concern globally, leading to significant social and medical expenses. Magnetic Resonance Imaging (MRI) is widely used for evaluating lumbar intervertebral disc degeneration due to its non-invasive nature and superior tissue differentiation capabilities. However, traditional 2D image analysis methods are often hindered by noise and various external factors, complicating accurate diagnosis and surgical planning. To address these challenges, this study investigates the application of deep learning models for lumbar vertebrae segmentation in MRI images. We employ U-Net, HRNet, and EfficientNet architectures, to develop an accurate segmentation model. The U-Net model, characterized by its encoding and decoding phases, demonstrated superior performance with a Precision of 0.9809, Recall of 0.9715, F1-score of 0.9742, and mAP of 0.7084. Comparatively, HRNet and EfficientNet also showed promising results, with HRNet achieving a Precision of 0.6684, Recall of 0.9153, F1-score of 0.77208, and mAP of 0.6568, while EfficientNet achieved a Precision of 0.7995, Recall of 0.9247, F1-score of 0.8491, and mAP of 0.7666. Our findings indicate that deep learning models, particularly U-Net, can significantly enhance the accuracy and efficiency of lumbar vertebrae segmentation in MRI images. This advancement holds potential for improving clinical diagnostics and surgical planning. Future work will focus on refining these models with larger datasets and exploring additional architectures to further enhance segmentation performance and robustness.**

***Keywords:* Medical Image Segmentation, U-Net, Convolutional Neural Network, Lumbar MRI, Performance Evaluation**

I. Introduction

Lower back pain is a prevalent health issue globally, resulting in substantial social and medical costs [1]. Magnetic resonance imaging (MRI), due to its non-invasive nature and ability to differentiate various tissues, is considered an ideal tool for evaluating lumbar intervertebral disc degeneration in clinical and research settings. Traditional detection methods in two-dimensional imaging are susceptible to noise, equipment, human, and environmental factors, significantly hindering physicians in diagnosing and planning surgeries based on the obtained images [2, 3]. Spinal image segmentation, which utilizes pixel features to separate pathological regions from muscle tissue, is crucial for analyzing spinal pathological structures and guiding spinal surgeries.

With the rapid development of artificial intelligence, convolutional neural networks (CNNs), known for their powerful nonlinear feature extraction capabilities, have greatly advanced research in image segmentation [4]. Deep learning based image segmentation methods can achieve excellent results in medical image segmentation. In terms of segmentation efficiency, deep learning have significant advantages over traditional machine learning algorithms [5]. This study aims to construct an accurate lumbar MRI segmentation model using CNN frameworks, specifically U-Net, HRNet, and EfficientNet architectures.

II. Related Work

Accurately segmenting disease areas and key tissues from various medical images provides a visual basis for clinical treatment, addressing visualization issues in many diagnoses. Traditional spinal image segmentation methods are mainly divided into semi-automatic and non-automatic categories [2]. Non-automatic segmentation methods can be completed based on differences in image thresholds and pixel values. Semi-automatic segmentation methods somewhat alleviate the operational pressure on physicians, but the segmentation speed still falls short of real-time segmentation requirements [6]. Although traditional medical image segmentation techniques have significantly advanced modern medical imaging, they struggle to meet the faster, more accurate, and more efficient demands of modern medical technology [7].

Deep learning-based image segmentation techniques have emerged in new forms. Long et al.[8] proposed Fully Convolutional Networks (FCN), which hold significant importance in the history of image semantic segmentation development. FCN addresses many

Pattern Recognition and Information Processing (PRIP2025) :
Proceedings of the 17th International Conference, 16–18 Sept. 2025, Minsk, Belarus. – Minsk : UIIP NASB, 2025. – 496 p. – ISBN 978-985-7198-20-7.

bottlenecks in semantic segmentation, surpassing traditional advanced levels and providing end-to-end, pixel-to-pixel segmentation solutions. However, FCN, as the first-generation semantic segmentation model, has many areas for improvement. Ronneberger et al. [9] modified FCN's network structure to change network connection patterns, proposing the U-Net model that emphasizes the connection of image features during training. U-Net, with its lightweight structure and good training effect on small sample sets, has become a classic end-to-end method in the segmentation field (Fig. 1). Prior research indicates that U-Net demonstrates superior performance in MRI vertebral segmentation.

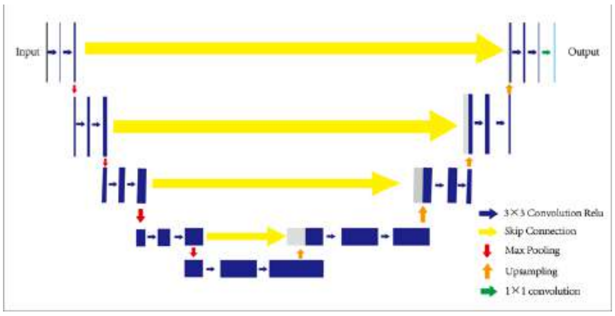

Fig. 1. U-Net Model

HR-Net [10] is a high-resolution network designed to achieve precise feature extraction and classification by maintaining high-resolution feature maps. The core design principle of HR-Net involves processing feature maps at different resolutions through multiple parallel branches, with feature exchange occurring at each stage of the network to preserve high-resolution information. The main structure of HR-Net comprises four stages, each containing multiple parallel convolutional modules that process feature maps of varying resolutions, which are then fused through feature exchange layers. The key feature is the maintenance of high-resolution feature maps to enable accurate feature extraction and classification (Fig. 2). Previous studies have shown that HR-Net achieves excellent results in human keypoint detection tasks, effectively identifying key body parts [10–12].

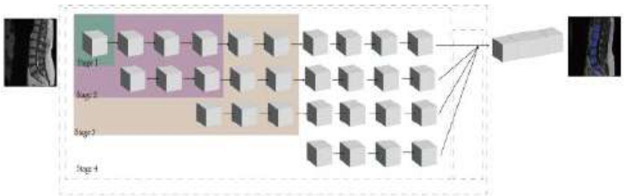

Fig. 2. HR-Net Model

EfficientNet [13] is an efficient CNN model that achieves high performance with reduced parameters and computational cost through comprehensive scaling of network depth, width, and resolution, thereby enabling efficient feature extraction and classification. EfficientNet-B0, the smallest version of EfficientNet, boasts fewer parameters and lower computational requirements. The main structure of EfficientNet-B0 consists of two convolutional modules, sixteen MBConv modules, a global average pooling layer, a dropout layer, and a fully connected layer, with the core being the sixteen MBConv modules, which achieve efficient feature extraction through the continuous stacking of multiple MBConv modules (Fig. 3). Existing research suggests that EfficientNet can effectively identify vertebrae and intervertebral discs in vertebral MRI [14].

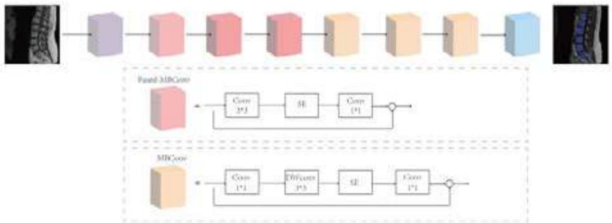

Fig. 3. EfficientNet model

However, the relative performance of the U-Net Model, HR-Net Model, and EfficientNet model in MRI vertebral segmentation, and which algorithm offers the optimal application effect, remains a subject of ongoing debate and lacks a definitive consensus.

III. Proposed Apporach

Our segmentation method is based on the convolutional neural networks, utilizing U-Net, HRNet, and EfficientNet architectures. The dataset comprises a total of 514 lumbar spine MRI images, which were divided into training, testing, and validation sets according to a distribution ratio of 70 %, 15 %, and 15 %, respectively.

A. Image Dataset

The training, testing, and validation datasets contain 514 lumbar MRI images with various lumbar diseases. The lumbar vertebrae mainly include the L1-L5 and S1 regions. All images were manually annotated and verified by experts, outlining the vertebrae contours.The partial images of the dataset are shown in Fig. 4.

Fig. 4. The partial images of the dataset

Pattern Recognition and Information Processing (PRIP2025) :
Proceedings of the 17th International Conference, 16–18 Sept. 2025, Minsk, Belarus. – Minsk : UIIP NASB, 2025. – 496 p. – ISBN 978-985-7198-20-7.

B. Experimental environment

The experimental environment for Comparative Analysis of Deep Learning Models for Lumbar Vertebrae Segmentation in MRI Images is shown in Table.

EXPERIMENTAL ENVIRONMENT

Software and hardware names	Specific parameters
CPU	Intel(R) Xeon(R) Gold 6226R
Memory	64GB
GPU	Quadro RTX 5000
Operating system	Window 10 Professional Edition
Software environment	Python3.12.7+torch2.6.0+cu118

C. Results and Discussion

In our experiments, we concentrated on precise vertebrae segmentation in lumbar MRI images. The U-Net model employed in this study features an encoding phase consisting of convolution operations and max-pooling layers, while the decoding phase utilizes convolution operations and deconvolution layers, culminating in a Softmax layer for classification. We also incorporated HRNet and EfficientNet models for comparative analysis. Following the input of lumbar MRI image data into the U-Net, HRNet, and EfficientNet models for segmentation prediction, the resulting segmentation masks are illustrated in Fig. 5-7, respectively.

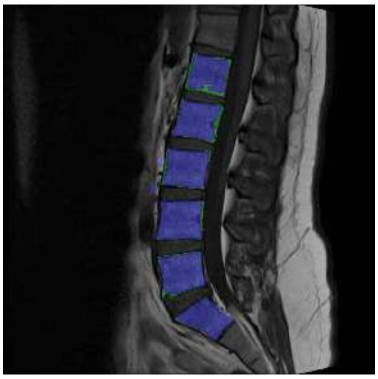

Fig. 5. Vertebrae segmentation by U-Net model

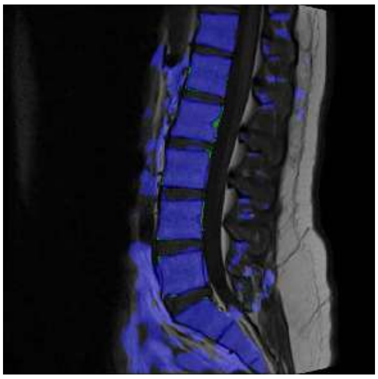

Fig. 6. Vertebrae segmentation by HRNet model

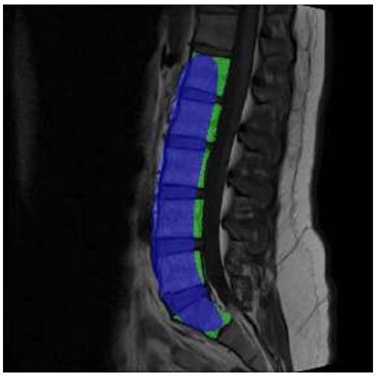

Fig. 7. Vertebrae segmentation by Efficient Net model

Quantitative evaluation metrics for each model are summarized as follows: U-Net achieved a Precision of 0.9809, Recall of 0.9715, F1-score of 0.9742, and mAP of 0.7084. HRNet obtained a Precision of 0.6684, Recall of 0.9153, F1-score of 0.77208, and mAP of 0.6568. EfficientNet yielded a Precision of 0.7995, Recall of 0.9247, F1-score of 0.8491, and mAP of 0.7666.As shown in Table II.

TABLE I. ERFORMANCE OF DIFFERENT MODELS

	U-Net	HRNet	Efficient Net
Precision	0.9809	0.6684	0.7995
Recall	0.9715	0.9153	0.9247
F1 Score	0.9742	0.7208	0.8491
mAP	0.7084	0.6568	0.7666

IV. CONCLUSION

This study explored the application of deep learning models, specifically U-Net, HRNet, and EfficientNet, for precise vertebrae segmentation in lumbar MRI images. Our results demonstrate the feasibility of these architectures for this task. While all models showed promise, U-Net achieved the highest overall performance in terms of Precision, Recall, F1-score, and mAP. This suggests its suitability for automated lumbar vertebrae segmentation.

REFERENCE

[1] R. A. Deyo, J. N. Weinstein, "Low Back Pain", N. Engl. J. Med., 2001, vol. 344(5), pp. 363-70.

[2] P. Yang, W. Song, X. Zhao, et al., "An Improved Otsu Threshold Segmentation Algorithm", International Journal of Computational Science and Engineering, 2020, vol. 22(1), pp. 146-53.

[3] K. Zhang, W. Lu, P. Marziliano, "Automatic Knee Cartilage Segmentation From Multi-Contrast Mr Images Using Support Vector Machine Classification With Spatial Dependencies", J. Magn. Reson. Imaging, 2013, vol. 31(10), pp. 1731-43.

[4] S. He, Q. Li, X. Li, M. Zhang, "A Lightweight Convolutional Neural Network Based on Dynamic Level-Set Loss Function ror Spine Mr Image Segmentation", J. Magn. Reson. Imaging, 2024, vol. 59(4), pp. 1438-53.

Pattern Recognition and Information Processing (PRIP2025) :
Proceedings of the 17th International Conference, 16–18 Sept. 2025, Minsk, Belarus. – Minsk : UIIP NASB, 2025. – 496 p. – ISBN 978-985-7198-20-7.

[5] C. Iriondo, V. Pedoia, S. Majumdar, "Lumbar Intervertebral Disc Characterization Through Quantitative Mri Analysis: An Automatic Voxel-Based Relaxometry Approach", J. Magn. Reson. Med., 2020, vol. 84(3), pp. 1376-90.

[6] L. Bi, Y. Peng, D. D. Feng, J. Kim, "Semi-Mamba: Improving Medical Image Segmentation Via Semi-Automatic Mamba", Network, Cham, F, 2025, Springer Nature Switzerland.

[7] T. Badriyah, N. Sakinah, I. Syarif, D. R. Syarif, "Segmentation Stroke Objects Based on Ct Scan Image Using Thresholding Method"; Proceedings of The 2019 First International Conference on Smart Technology & Urban Development (Stud), F 13-14, Dec. 2019.

[8] E. Shelhamer, J. Long, T. Darrell, "Fully Convolutional Networks For Semantic Segmentation", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39(4), 2017, pp. 640-51.

[9] O. Ronneberger, P. Fischer, T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation", Cham, F, 2015, Springer International Publishing.

[10] J. I. T. Noel, J. R. A., G. Sushil, N. Ruban, "Performance Analysis of Lightweight Cnn Models to Segment Infectious Lung Tissues of Covid-19 Cases from Tomographic Images", Peerj Computer Science, vol. 7, 2021, pp. E368-E.

[11] X. Dai, X. Ye, J. Ren, et al., "Construction of the Preoperative Staging Prediction Model for Cervical Cancer Based on Deep Learning and Mri: A Retrospective Study", Frontiers in Oncology, vol. 15, 2025, pp. 1557486.

[12] H. Marcadrien, C O A, P M C, et al., "The Impact of Fatty Infiltration on Mri Segmentation of Lower Limb Muscles in Neuromuscular Diseases: A Comparative Study of Deep Learning Approaches", Journal of Magnetic Resonance Imaging, vol. 58(6), 2023, pp. 1826-35.

[13] B. S. S. Dias, R. Querrer, P. T. Figueiredo, et al., "Osteoporosis Screening: Leveraging Efficientnet with Complete and Cropped Facial Panoramic Radiography Imaging", Biomedical Signal Processing and Control, vol. 100(Pb), 2025, pp. 107031.

[14] K. Kurachka, U. Kamrakou, N. Masalitina, "The Automated Classification System for Lumbar Spine Anatomic Elements", Nonlinear Dynamics And Applications, vol. 23, 2017, pp. 127-134.

Pattern Recognition and Information Processing (PRIP2025) :
Proceedings of the 17th International Conference, 16–18 Sept. 2025, Minsk, Belarus. – Minsk : UIIP NASB, 2025. – 496 p. – ISBN 978-985-7198-20-7.