Application of Deep Learning to Crack Segmentation in Neutron CT Images of Ancient Writing Knives

Authors: Li, Mr. Jiancong, wang, Dr. Shengxiang, Shu, Mr. Xin, Dong, Le, Dr. Yong Lei, Chen, Prof. Jie, wang, Dr. Shengxiang

Date: 2025-08-03T15:10:06+00:00

Abstract

Neutron CT imaging offers unique advantages in metal defect detection and cultural relic analysis, particularly for exploring internal structures of ancient artifacts like writing knives, due to its high penetrability and hydrogen sensitivity. Accurate segmentation of its images is critical for defect detection, with crack segmentation in writing knife images being key to understanding craftsmanship and preservation. However, complex structures in these images—neutron scattering noise, blurred multi-material interfaces and overlapping gray scales hinder precise crack segmentation. Traditional algorithms, reliant on manual tuning and single-feature extraction, lack accuracy: they roughly distinguish macroscopic structures but fail to segment fine cracks in blade edges. This study addresses this by applying deep learning to crack segmentation in writing knife neutron CT images, using BSEResU-Net (a residual U-Net with SE attention). Trained on a small manually annotated dataset of The Western Han writing knife from China Spallation Neutron Source (CSNS), the model was validated on full-knife crack segmentation. Results show its superiority, it obtained an Area Under the ROC Curve (AUC) of 0.9793 and an F1 score of 0.9089 on the dataset, accurately capturing fine cracks. Compared with threshold segmentation, about 70% more cracks can be segmented. This framework resolves neutron data scarcity, provides an innovative solution for cultural heritage defect detection, and advances deep learning in multimodal penetrating imaging.

Full Text

Preamble

Application of Deep Learning to Crack Segmentation in Neutron CT Images of Ancient Writing Knives

Jiancong Li[1],[2], Shengxiang Wang[2],[3],†, Xin Shu[2],[3], Le Dong[2],[3], Yong Lei ,‡, and Jie Chen[2],[3],§

[1]School of Computer Science and Technology, Dongguan University of Technology, Dongguan 523803, China

[2]Spallation Neutron Source Science Center (SNSSC), Dongguan 523803, China

[3]Institute of High Energy Physics, Chinese Academy of Sciences (CAS), Beijing 100049, China

 The Palace Museum, No. 4 Jingshan Front Street, Dongcheng District, Beijing 100009, China

Neutron CT imaging offers unique advantages in metal defect detection and cultural heritage analysis, particularly for exploring the internal structures of ancient artifacts such as writing knives, due to its high penetrability and hydrogen sensitivity. Accurate segmentation of these images is critical for defect detection, with crack segmentation in writing knife images being key to understanding craftsmanship and preservation states. However, complex structures in these images—including neutron scattering noise, blurred multi-material interfaces, and overlapping gray scales—hinder precise crack segmentation. Traditional algorithms, which rely on manual tuning and single-feature extraction, lack sufficient accuracy: they can roughly distinguish macroscopic structures but fail to segment fine cracks at blade edges.

This study addresses these challenges by applying deep learning to crack segmentation in writing knife neutron CT images, utilizing BSEResU-Net (a residual U-Net with SE attention). Trained on a small manually annotated dataset of a Western Han writing knife from the China Spallation Neutron Source (CSNS), the model was validated on full-knife crack segmentation. Results demonstrate its superiority, achieving an Area Under the ROC Curve (AUC) of 0.9793 and an F1 score of 0.9089 on the dataset, accurately capturing fine cracks. Compared with threshold segmentation, approximately 70% more cracks can be segmented. This framework resolves neutron data scarcity, provides an innovative solution for cultural heritage defect detection, and advances deep learning in multimodal penetrating imaging.

Keywords: Neutron CT; ancient writing knife; crack segmentation; BSEResU-Net.

Introduction

The advancement of imaging technologies has propelled progress across multidisciplinary fields. Among them, neutron imaging, as a typical penetrating

detection technique, demonstrates irreplaceable application value in materials science, cultural heritage conservation, and other domains by virtue of its unique physical properties [**?**, **?**]. Neutrons, with their strong penetrability and sensitivity to hydrogen, enable non-destructive testing for internal defects in metal materials—precisely locating micro-cracks, air holes, and other hidden hazards in metal components. They also facilitate non-invasive analysis in cultural relic research, such as revealing the casting techniques of ancient bronze wares or the stratification of mural pigments.

Image segmentation [**?**] is a pivotal step in the data processing pipeline of neutron imaging, primarily serving to partition raw images into regions with physical significance. This process not only enables effective separation of target objects from the background but also provides the data foundation for subsequent applications such as defect identification and compositional analysis. In practical scenarios, the accuracy of segmentation directly determines the reliability of analysis results, including the diagnosis of internal material defects and the structural interpretation of cultural relics. As a crucial link connecting imaging data acquisition and in-depth analysis, improving the performance of image segmentation technology is of great significance for advancing penetrating detection techniques. Traditional segmentation algorithms, such as thresholding [**?**] and edge detection [**?**], despite their simplicity in principle, exhibit limitations when confronted with typical challenges of these imaging modalities. In neutron images, low-contrast targets, complex background noise, and regions with mixed materials [**?**, **?**] are common. These challenges, coupled with the reliance of traditional algorithms on manual parameter tuning and single-feature extraction, often lead to problems such as missed targets and blurred boundaries in segmentation results, making it difficult to meet the requirements of high-precision analysis.

In recent years, breakthroughs in deep learning for image segmentation [**?**] have offered new solutions to these technical challenges. Convolutional neural networks (CNNs) [**?**], through multi-layer feature extraction and non-linear mapping, can automatically learn high-order representations of hydrogen distribution and defects in neutron images, effectively enhancing segmentation robustness across different imaging modalities. Long et al. proposed the Fully Convolutional Network (FCN) [**?**], a milestone in deep learning-based semantic image segmentation models. The core contribution of FCN lies in its ability to accept input images of arbitrary sizes and output corresponding segmentation maps, achieved by converting fully connected layers in traditional CNNs into convolutional layers. The advent of FCN marked a significant advancement in deep learning for image segmentation. Building on this, Ronneberger et al. developed U-Net [**?**], a classic architecture specifically designed for medical image segmentation. The core structure employs a symmetric encoder-decoder design: the contracting path (downsampling) extracts contextual features via convolution and max-pooling, while the expansive path (upsampling) restores resolution using transposed convolutions, with feature fusion enabled by skip connections. This design allows the network to capture global semantic information while

precisely localizing target regions—the downsampling path compresses features layer-by-layer to enhance abstract representation, the upsampling path restores spatial dimensions via transposed convolutions, and skip connections directly transmit original features from the downsampling path to preserve details otherwise lost during pooling. U-Net excels in scenarios with scarce medical data, improving generalization through data augmentation such as rotation and scaling to learn effective feature representations from minimal samples, and remains a foundational architecture in medical segmentation to date.

The Energy Resolved Neutron Imaging Spectrometer (ERNI) of CSNS [**?**] has carried out neutron CT detection on an ancient writing knife collected by the Palace Museum. To deeply analyze its internal structure, the research team performed image segmentation work. However, traditional image segmentation methods can only roughly distinguish different structures inside the writing knife and struggle to achieve accurate segmentation of subtle cracks in the blade. To address this technical bottleneck, deep learning methods are introduced in this study to provide a more efficient solution to these problems. The raw data undergo denoising and preprocessing, followed by 3D reconstruction and annotation to establish a specialized neutron image dataset. By integrating the BSEResU-Net [**?**] segmentation framework—an architecture that fuses residual learning and SE attention mechanism into the traditional U-Net—this research conducts image segmentation on real-world datasets, significantly enhancing feature extraction efficiency. The findings not only provide technical support for scientific challenges such as precise material defect identification and fine-scale structural analysis of archaeological relics but also bridge the data and methodological gaps in deep learning applications for neutron imaging, fostering technological integration across interdisciplinary research.

II. Materials and Methods

BSEResU-Net, as an innovative variant within the U-Net family, has systematically reconstructed and upgraded the classic architecture. It groundbreakingly fuses the SE attention mechanism [**?**] with residual connections [**?**] in-depth, building an efficient framework better adapted to complex feature processing. As shown in the architecture diagram (Fig. 1 [Figure 1: see original paper]), it strengthens gradient transmission and suppresses network degradation through residual operations, leverages the SE attention mechanism to precisely calibrate channel feature weights, and cooperates with components such as skip connections and transpose convolutional layers to form a complete encoding-decoding process.

In terms of performance, this network demonstrates two remarkable advantages. First, regarding resolution compatibility, its flexible hierarchical design allows seamless adaptation to multi-source heterogeneous data with different resolutions, accurately parsing feature information from macro scenes to micro details. Second, concerning architectural flexibility, it supports adjusting the number of residual blocks and network depth as needed, achieving a dynamic balance be-

tween model complexity (computational cost) and segmentation performance. This customizable feature, combined with the synergistic benefits of residual connections and the attention mechanism, enables BSEResU-Net to deeply tap into the potential value of data and strengthen the expression of key features. Consequently, in complex scene segmentation tasks, it achieves a dual breakthrough in precision and robustness, effectively coping with challenges such as noise interference and target blurriness, and providing a superior technical solution for fields like medical image analysis and remote sensing image interpretation.

A. Residual Convolution

In the realm of deep learning, as the number of network layers grows, linearly stacked architectures often encounter the formidable challenges of vanishing gradients [?] and performance degradation. These issues severely impede the training of deep neural networks, limiting their ability to extract hierarchical features effectively. Residual connections emerge as a groundbreaking solution, revolutionizing network design through an innovative skip connection [?] mechanism. As shown in Fig. 2 Figure 2: see original paper, residual connections establish direct "shortcut paths" that bypass intermediate layers, enabling seamless transmission of low-level feature information from early network stages to deeper layers. This architectural innovation allows the network to focus on learning the residual differences between the input and desired output, rather than reconstructing the entire mapping from scratch. By doing so, it not only mitigates the exponential decay of gradients during backpropagation [?] but also reduces the computational burden on individual layers, facilitating more efficient learning.

The unique design of residual connections brings dual advantages: preserving gradient integrity throughout the network hierarchy and accelerating the optimization process. Empirical studies have demonstrated that this approach significantly enhances training stability, enabling faster convergence and preventing performance degradation even in extremely deep architectures. As a result, residual connections have become a fundamental building block for developing state-of-the-art neural networks, empowering researchers to construct highly complex yet robust models capable of achieving superior performance across diverse applications.

B. SE Attention Mechanism

The SE Attention Mechanism represents a highly influential innovation in deep learning that has revolutionized feature learning in Convolutional Neural Networks (CNNs) with its unique channel-wise adaptive adjustment strategy. Rooted in the core concept of "squeeze-and-excitation," this mechanism pioneers a novel computational paradigm for channel attention, enabling precise discrimination of the significance of different channel features in data and dynamic optimization of feature representation.

The SE mechanism refines channel features through a two-step process, as shown in Fig. 2 Figure 2: see original paper. First, the Squeeze operation condenses the spatial information of each channel into a global feature descriptor using techniques like global average pooling [?]. This effectively captures the global context within channels, abstracting feature maps into comprehensive representations. Subsequently, the Excitation operation leverages a Multi-Layer Perceptron (MLP) [?] to establish non-linear relationships across channels, adaptively generating weight coefficients for each channel. These coefficients quantify the contribution of individual channel features to the final output. By fusing them with the original feature maps, the model amplifies the expression of crucial channels while suppressing redundant or interfering information. This adaptive weight allocation endows the SE mechanism with exceptional capabilities in feature selection and enhancement.

C. Loss Function

In image segmentation tasks, a single loss function is often difficult to balance segmentation accuracy and class balance. The Dice loss function [?] focuses on measuring the overlap between predicted results and true labels, which can effectively deal with data imbalance problems. The cross-entropy loss function [?] optimizes category prediction accuracy from a probability perspective and performs well in capturing detailed features.

To leverage the advantages of both, this study uses a joint loss function combining Dice loss and cross-entropy loss, constructing an optimization objective that considers both pixel-level classification accuracy and regional integrity through linear weighting, thereby achieving a more balanced and accurate image segmentation effect. The specific expression of the joint loss function is as follows:

$$L = \alpha \cdot CEL + (1 - \alpha) \cdot DL$$

where CEL stands for cross-entropy loss function; DL stands for Dice loss function; and   is a hyperparameter controlling the contribution of cross-entropy loss and Dice loss. By adjusting , the optimal combination of loss functions can be found for different tasks to improve segmentation accuracy.

The cross-entropy loss function is often used in classification tasks, especially in binary or multiclass classification problems. The goal is to minimize the difference between predicted values and true labels. A major drawback of the cross-entropy loss function is that in cases of imbalanced data, it may cause the model to over-optimize for the large class (majority class) while ignoring the small class (minority class). Therefore, in some highly imbalanced segmentation tasks, cross-entropy loss may not effectively handle the segmentation problem between foreground and background. The formula is as follows:

$$CEL = -\sum [t_i \ln(p_i) + (1 - t_i) \ln(1 - p_i)]$$

where N is the number of samples (or pixels, in the case of image segmentation tasks). t_i is the true label of the ith example, taking the value 0 or 1 (e.g., foreground or background in a segmentation task). p_i is the predicted probability value of the ith sample.

The Dice loss function is specifically designed for image segmentation. It is based on the Dice coefficient, which measures the overlap between the predicted region and the true region, and performs well especially when dealing with imbalanced data. Its strength lies in superior performance when facing imbalanced data, as it balances the contributions of foreground and background by directly optimizing their overlap. In tasks such as medical image segmentation, the foreground (e.g., tumors or blood vessels) usually occupies a smaller portion of the image, and Dice loss handles this effectively. The formula is as follows:

$$DL = 1 - \frac{2\sum_{i=1}^{N} p_i t_i + \epsilon}{\sum_{i=1}^{N} p_i^2 + \sum_{i=1}^{N} t_i^2 + \epsilon}$$

where N is the number of samples (pixels); p_i is the predicted probability of the ith pixel; t_i is the true label of the ith pixel; and   is a small smoothing factor used to avoid division by zero errors.

D. Evaluation Metrics

To evaluate segmentation performance and construct a scientific and reasonable evaluation system, the following key evaluation indicators are introduced. TP, TN, FP, and FN represent true positive, true negative, false positive, and false negative examples of segmentation, respectively. The three evaluation metrics are described as follows:

Se (Sensitivity/Recall) focuses on the model's ability to identify positive samples. It measures the proportion of actual positive instances correctly predicted as positive by the model. A higher Se value indicates a lower probability of missing relevant positive samples, making it particularly crucial for applications requiring high sensitivity. Sp (Specificity) evaluates the model's proficiency in distinguishing negative samples. It calculates the ratio of actual negative instances accurately classified as negative, with higher values signifying better discrimination against false positives. This metric is essential for scenarios where avoiding false alarms is critical. ACC (Accuracy) provides an overall assessment of model performance by measuring the percentage of correctly classified samples (both positive and negative) out of the total dataset. However, in scenarios with class imbalance, ACC may not adequately reflect the model's true capabilities, necessitating complementary analysis with metrics like Se and Sp to ensure comprehensive evaluation.

Ideally, a model should achieve high scores in both Se and Sp. However, there is often an inherent trade-off between these metrics—enhancing one may inad-

vertently compromise the other, especially in scenarios with imbalanced data distributions or when adjusting decision thresholds.

To address the limitations of single-metric evaluation and enable a more comprehensive, balanced assessment of model performance, two composite metrics have been proposed: F1-Score and Matthews Correlation Coefficient (MCC). F1-Score, a harmonic mean of Precision and Recall, prioritizes balancing the accuracy of positive predictions (Precision) with their completeness (Recall). MCC, a correlation coefficient that integrates all four confusion matrix categories (TP, TN, FP, FN), is particularly well-suited for imbalanced datasets. By accounting for both foreground-background discrimination and class distribution dynamics, these metrics provide multi-dimensional validation, overcoming single-metric biases and offering a more robust evaluation framework. F1-Score and MCC are described as follows:

$$F1 = 2 \times \frac{Pr \times Se}{Pr + Se}$$

$$MCC = \frac{TP/N - S \times P}{\sqrt{P \times S \times (1 - S) \times (1 - P)}}$$

where N denotes the total number of pixels in the image. Pr = TP/(TP+FP) represents precision, measuring the proportion of predicted positive pixels that are actually correct. S = (TP + FN)/N is the foreground ratio, indicating the proportion of actual foreground pixels in the entire image. P = (TP + FP)/N is the predicted positive ratio, reflecting the proportion of pixels classified as foreground by the model.

III. Experiment and Results

A. Datasets

The model dataset used in the experiment is the neutron CT dataset of a writing knife from the Western Han Dynasty. The Western Han writing knife was tested for neutron tomography at a distance of approximately 27 meters from the ERNI instrument at CSNS. The pinhole was moved to 20 mm, and the neutron wavelength range was set between 0.5 and 4.6 Angstroms by the chopper. For neutron tomography, all projections were recorded through an optical system that included a 50-micron-thick ZnS/6LiF scintillation screen, a CCD camera (Andor, Oxford Instruments), and a Nikon photographic lens (Nikon, Japan). The camera array had a size of 2048×2048 pixels, with each pixel being 15 microns in size. The magnification of the optical lens was adjusted to approximately 2 times, and the imaging field of view was 5.5×5.5 square centimeters. The acquired projection data were reconstructed in 3D using the FBP algorithm [?] with parallel bundles. The dataset after 3D reconstruction contained 2594 cross-sectional data slices. The original cross-sectional data were cropped to

a size of 308×308 pixels, and 200 representative cross-sectional images of the knife edge were selected from different regions for annotation and training.

In this study, the labelme [**?**] tool was used to complete the data annotation work. As open-source annotation software, labelme provides flexible manual annotation functionality, supporting researchers to accurately outline and define the category of target areas in image data. Its open-source nature not only reduces tool usage costs but also allows users to expand functionality through custom scripts according to specific task requirements, providing high-quality labeled datasets for subsequent model training and ensuring the standardization and availability of experimental data.

B. Experiment Setup

Regarding experimental hardware configuration and training strategy, the study employed an NVIDIA GeForce RTX 4060 Laptop GPU for computation with the CUDA 12.7 computing framework featuring 8GB of video memory. It should be noted that this configuration is only appropriate for addressing the computational requirements of small-scale training scenarios. Given the focus of this paper on verification methods, small-scale training was utilized to demonstrate the effectiveness of the proposed approach. For large-scale data processing, it is imperative to upgrade the hardware to overcome limitations in video memory capacity and computational power.

During the data preprocessing phase, input images were divided into 48×48 pixel blocks to achieve a balance between memory utilization efficiency and preservation of local features. This design choice not only prevents video memory overload but also ensures accurate capture of local characteristics such as cracks and pores, thereby providing a robust data foundation for model training. The data was augmented by rotating 90°, 180°, and 270°, and a total of 40,000 patches were used for training.

Regarding model training strategy, 100 training epochs were established to ensure the model converges to a stable state through sufficient iterations, thus guaranteeing the reliability of segmentation accuracy. A dynamic learning rate adjustment strategy was implemented: initially set at 2×10^2 to accelerate early convergence, it was gradually attenuated to 5×10^3 as training progressed. Refined parameter tuning was performed during the near-convergence stage of the model to balance convergence speed and accuracy. The optimizer employed stochastic gradient descent (SGD) [**?**], leveraging its efficient gradient update characteristics to drive iterative optimization of model parameters. This approach ultimately achieved high-fidelity alignment between segmentation results and real labels in tasks such as pearl inner wall segmentation and crack detail segmentation.

C. Results on Western Han Dynasty Writing Knife Neutron CT Dataset

From the experimental results presented in Fig. 3 [Figure 3: see original paper], it is evident that the BSEResU-Net model utilized in this study demonstrates superior performance in crack detail segmentation tasks. The proposed model effectively captures fine-scale features of crack images with significant variations in gray value distribution, achieving high-precision segmentation outcomes.

In contrast to traditional threshold-based segmentation methods, which can only provide rough approximations of target edge contours, BSEResU-Net leverages its deep residual structure and attention mechanism to accurately identify and segment complex crack morphologies within the writing knife. This approach effectively mitigates issues related to edge blurring and detail loss. Compared to classical models such as SegNet and U-Net, the proposed model exhibits distinct advantages in segmentation accuracy, edge localization precision, and detail preservation. When processing crack images characterized by multi-scale and multi-texture features, BSEResU-Net achieves more precise pixel-level classification of crack regions through its optimized feature extraction module and multi-level feature fusion strategy, thereby significantly enhancing the refinement of segmentation results.

According to the multi-dimensional evaluation metrics in Table 1 , the proposed BSEResU-Net method attains outstanding performance across all indicators, surpassing both SegNet and U-Net in every respect. Notably, the AUC metric reaches 0.9793, a value that leaves comparison models far behind and underscores the method' s superior capacity to discriminate between segmentation results and actual ground truth. While the Sp index of BSEResU-Net, at 0.9872, is marginally lower than SegNet' s 0.9898, overall BSEResU-Net still exhibits more robust comprehensive segmentation performance, especially demonstrating higher precision when handling detailed components.

Table 1. Performance comparison of different segmentation methods on the Western Han Dynasty writing knife neutron CT dataset

Methods	AUC↑	ACC↑	F1↑	MCC↑	Sp↑
SegNet					0.9898
U-Net					
BSEResU-Net	0.9793	0.9536	0.8621	0.9872	0.9089

The threshold segmentation result in Fig. 4 Figure 4: see original paper is limited by the single attribute of pixel gray value, which not only leads to blurred edges of cracks but also causes loss of structural information, making it difficult to clearly reveal subtle defects inside the writing knife. BSEResU-Net uses a multi-layer convolutional neural network to perform feature extraction and semantic segmentation of writing knife images. The segmentation results

in Fig. 4 show that with powerful feature learning capability, the proposed algorithm can not only accurately capture the fine boundaries of writing knife cracks and completely separate them from the complex background, but also deeply analyze the internal hierarchical structure of the blade, clearly revealing the direction, depth, and spatial relationship between cracks and the blade matrix. Compared with threshold segmentation results, the 3D model output by BSEResU-Net restores the actual form of the writing knife more realistically and depicts crack details with exquisite precision, providing more reliable data support for subsequent quality assessment and structural damage analysis of the writing knife. This fully demonstrates the technical superiority of the algorithm in the field of 3D structure segmentation of complex objects. Simultaneously, compared with segmentation results from U-Net and SegNet, it is found that due to insufficient segmentation by U-Net and SegNet on some slices, the final overall 3D rendering is partially missing, which further highlights the advantage of BSEResU-Net in segmentation effectiveness.

Although the BSEResU-Net proposed in this study demonstrated outstanding performance in CT image segmentation tasks, thorough analysis of experimental results revealed a typical area for optimization. In terms of segmentation accuracy, the quality of labeled data has an undeniable impact on model training effectiveness. During training set annotation, subjectivity in manual labeling and difficulty in locating boundaries of small structures led to widespread labeling errors, with some annotated boundaries deviating from actual target edges. These labeling errors are learned by the model during training, leading to over-segmentation in complex texture regions where the model misjudges texture-similar background regions as targets. Simultaneously, segmentation accuracy for small structures (such as millimeter-level cracks and sub-voxel-level pores) is insufficient. This not only reflects the model' s inadequate representation capability for fine-grained features under small-sample training conditions but also highlights how labeled data accuracy constrains model learning effectiveness.

IV. Conclusion

This study applies a novel deep learning framework, BSEResU-Net, which deeply integrates the residual U-Net architecture with the SE attention mechanism to construct an image segmentation model with efficient feature extraction and semantic information enhancement capabilities. To systematically evaluate model performance, a small-scale dataset for neutron CT was constructed. Through rigorous cross-validation training and testing, experimental results demonstrate that BSEResU-Net exhibits outstanding segmentation performance on the neutron CT dataset, capable of accurately capturing subtle structural differences and boundaries in images. Compared with current mainstream image segmentation algorithms, this model achieves improvements in key evaluation metrics, effectively overcoming technical bottlenecks of traditional threshold segmentation methods such as insufficient accuracy and poor adaptability, providing an innovative solution for high-precision semantic segmentation of neutron CT

images.

Acknowledgements

References

[1] Ziesche, Ralf F., et al. "4D imaging of lithium-batteries using correlative neutron and X-ray tomography with a virtual unrolling technique." *Nature communications* 11.1 (2020): 777. DOI: 10.1038/s41467-019-13943-3

[2] Lehmann, Eberhard H., et al. "The XTRA option at the NEUTRA facility —more than 10 years of bi-modal neutron and X-ray imaging at PSI." *Applied Sciences* 11.9 (2021): 3825. DOI: 10.3390/app11093825

[3] Yu, Ying, et al. "Techniques and challenges of image segmentation: A review." *Electronics* 12.5 (2023): 1199. DOI: 10.3390/electronics12051199

[4] Roy, Payel, et al. "Adaptive thresholding: A comparative study." 2014 International conference on control, Instrumentation, communication and Computational Technologies (ICCICCT). IEEE, 2014. DOI: 10.1109/ICCICCT.2014.6993140

[5] Sun, Rui, et al. "Survey of image edge detection." *Frontiers in Signal Processing* 2 (2022): 826967. DOI: 10.3389/frsip.2022.826967

[6] Rouchon, Amélie, Andrea Zoia, and Richard Sanchez. "A new Monte Carlo method for neutron noise calculations in the frequency domain." *Annals of Nuclear Energy* 102 (2017): 465-475. DOI: 10.1016/j.anucene.2016.11.035

[7] Meganck, Jeffrey A., et al. "Beam hardening artifacts in micro-computed tomography scanning can be reduced by X-ray beam filtration and the resulting images can be used to accurately measure BMD." *Bone* 45.6 (2009): 1104-1116. DOI: 10.1016/j.bone.2009.07.078

[8] Minaee, Shervin, et al. "Image segmentation using deep learning: A survey." *IEEE transactions on pattern analysis and machine intelligence* 44.7 (2021): 3523-3542. DOI: 10.1109/TPAMI.2021.3059968

[9] LeCun, Yann, et al. "Backpropagation applied to handwritten zip code recognition." *Neural computation* 1.4 (1989): 541-551. DOI: 10.1162/neco.1989.1.4.541

[10] Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." *Proceedings of the IEEE conference on computer vision and pattern recognition.* 2015. DOI: 10.1109/cvpr.2015.7298965

[11] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." *Medical image computing*

and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. Springer international publishing, 2015. DOI: 10.1007/978-3-319-24574-4_28

[12] Chen, Jie, et al. "The energy-resolved neutron imaging instrument at the China spallation neutron source." *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1064 (2024): 169460. DOI: 10.1016/j.nima.2024.169460

[13] Li, Di, and Susanto Rahardja. "BSEResU-Net: before-activation residual U-Net with squeeze-and-excitation attention-based retinal vessel segmentation." *Computer Methods and Programs in Biomedicine* 205 (2021): 106070. DOI: 10.1016/j.cmpb.2021.106070

[14] Hu, Jie, Li Shen, and Gang Sun. "Squeeze-and-excitation networks." *Proceedings of the IEEE conference on computer vision and pattern recognition.* 2018. DOI: 10.1109/cvpr.2018.00745

[15] Szegedy, Christian, et al. "Inception-v4, inception-resnet and the impact of residual connections on learning." *Proceedings of the AAAI conference on artificial intelligence.* Vol. 31. No. 1. 2017. DOI: 10.1609/aaai.v31i1.11231

[16] Hochreiter, Sepp. "The vanishing gradient problem during learning recurrent neural nets and problem solutions." *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 6.02 (1998): 107-116. DOI: 10.1142/S0218488598000094

[17] Zhou, Zongwei, et al. "Unet++: Redesigning skip connections to exploit multiscale features in image segmentation." *IEEE transactions on medical imaging* 39.6 (2019): 1856-1867. DOI: 10.1109/TMI.2019.2959609

[18] Amari, Shun-ichi. "Backpropagation and stochastic gradient descent method." *Neurocomputing* 5.4-5 (1993): 185-196. DOI: 10.1016/0925-2312(93)90006-O

[19] Kumar, R. Lokesh, et al. "Multi-class brain tumor classification using residual network and global average pooling." *Multimedia Tools and Applications* 80.9 (2021): 13429-13438. DOI: 10.1007/s11042-020-10335-4

[20] Kruse, Rudolf, et al. "Multi-layer perceptrons." *Computational intelligence: a methodological introduction.* Cham: Springer International Publishing, 2022. 53-124.

[21] Sudre, Carole H., et al. "Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations." *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, Proceedings 3. Springer International Publishing, 2017.

[22] Ho, Yaoshiang, and Samuel Wookey. "The real-world-weight cross-entropy loss function: Modeling the costs of mislabeling." *IEEE access* 8 (2019): 4806-4813. DOI: 10.1109/ACCESS.2019.2962617

[23] Scherl, Holger, et al. "Implementation of the FDK algorithm for cone-beam CT on the cell broadband engine architecture." *Medical Imaging 2007: Physics of Medical Imaging.* Vol. 6510. SPIE, 2007. DOI: 10.1117/12.708754

[24] Willemink, Martin J., and Peter B. Noël. "The evolution of image reconstruction for CT—from filtered back projection to artificial intelligence."*European radiology* 29 (2019): 2185-2195. DOI: 10.1007/s00330-018-5810-7

[25] Russell, Bryan C., et al. "LabelMe: a database and web-based tool for image annotation." *International journal of computer vision* 77 (2008): 157-173. DOI: 10.1007/s11263-007-0090-8

[26] Bottou, Léon. "Stochastic gradient descent tricks." *Neural networks: tricks of the trade: second edition.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. 421-436.

[27] Wang, Shengxiang, et al. "Design of neutron and x-ray CT system in ERNI of CSNS."*Advanced Optical Manufacturing Technologies and Applications 2022; and 2nd International Forum of Young Scientists on Advanced Optical Manufacturing (AOMTA and YSAOM 2022).* Vol. 12507. SPIE, 2023. DOI: 10.1117/12.2656869

[28] Wang, Sheng-Xiang, et al. "Ring artifacts correction based on the projection-field in neutron CT."*Chinese Physics B* 30.5 (2021): 050601. DOI: 10.1088/1674-1056/abd743

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.