Journal of Imaging Informatics in Medicine
https://doi.org/10.1007/s10278-026-01890-7

A Landmark-Guided Dual-Stream Synergistic Framework for Automated Intracranial Aneurysm Detection in Magnetic Resonance Angiography

Doyeon Kim[1] · Jieun Park[1] · Hyeonsik Yang[1] · Gi-youn Kim[1] · Jiyeon Lee[1] · Donghyeon Kim[1] · Hyun Jin Han[2] · Keun Young Park[2] · Minho Lee[1]

Received: 8 September 2025 / Revised: 13 February 2026 / Accepted: 18 February 2026

Abstract

Early and accurate detection of intracranial aneurysms (IAs) is critical for preventing rupture; however, manual interpretation of time-of-flight magnetic resonance angiography (TOF-MRA) scans requires time-intensive review, increasing clinician workload. Although deep-learning methods offer promising solutions, existing approaches often rely on impractical vessel segmentation or inefficient patch sampling. Herein, we propose a dual-stream synergistic framework for IA detection and segmentation that balances diagnostic accuracy with clinical feasibility. The first stage performs landmark-guided candidate detection, beginning with a U-Net-based model trained using a novel adaptive loss function to predict the locations of 18 vascular landmarks. These landmarks guide the extraction of coordinate-aware patches, which are then analyzed using a hybrid UNETR-FPN model to identify potential aneurysm candidates. The second stream employs a fine-grained segmentation model (nnU-Net with a residual encoder) operating on the full MRA volume to delineate lesion boundaries. Final predictions are generated through conditional fusion of both streams, prioritizing candidate detection and refining shapes with segmentation information. The framework was trained and evaluated on 1055 TOF-MRA scans, including two distinct internal test sets. The landmark localization model achieved >0.97 sensitivity in capturing the ground-truth aneurysms within generated patches. On the two test sets, the framework achieved lesion-wise sensitivities of 0.87 and 0.82 with corresponding false-positive rates of 1.23 and 1.17 per case. Performance was robust across anatomical locations but declined for aneurysms ≤ 3 mm. This landmark-guided dual-stream framework provides strong performance for IA detection while reducing annotation demands, offering a clinically practical tool for cerebrovascular diagnostics.

Keywords Intracranial aneurysm · Deep learning · Magnetic resonance angiography (MRA) · Landmark detection · Image segmentation · Computer-aided detection

Introduction

Artificial Intelligence (AI) has been increasingly adopted across diverse medical domains and is now widely used as a second reader or clinical decision-support tool, assisting clinicians in tasks ranging from classification and diagnosis to treatment selection. Recent studies have demonstrated that AI- and machine learning-based systems are increasingly integrated into medical imaging workflows, assisting with tasks ranging from lesion detection and classification to treatment planning and outcome prediction, thereby underscoring their broad applicability in healthcare. This trend is equally evident in neuroimaging. Given the anatomical variability and structural complexity of the cerebral vasculature, automated assistance is particularly valuable for intracranial aneurysm (IA) detection.

IA is a localized dilation of the weak walls of the cerebral arteries [1]. Approximately 3% of healthy adults have intracranial aneurysms, with approximately 1% of these patients at risk for rupture [2, 3]. Most IAs remain asymptomatic until rupture, which can lead to life-threatening subarachnoid hemorrhage, with corresponding

✉ Minho Lee
minho.lee@neurophet.com

[1] Research Institute, Neurophet Inc., 12 F, 124, Teheran-Ro, Gangnam, Seoul, Republic of Korea

[2] Department of Neurosurgery, Severance Hospital, Yonsei University College of Medicine, Seoul, Republic of Korea

Published online: 18 May 2026

high morbidity and mortality rates. Therefore, early IA detection is crucial for timely medical interventions and prevention.

IAs are most commonly detected using digital subtraction angiography (DSA) and magnetic resonance angiography (MRA), the primary neuroimaging modalities widely used in clinical practice. Among these modalities, DSA is considered the gold standard for IA diagnosis because of its clear geometrical representation of aneurysm size and shape. Despite these advantages, its routine use is limited because of its invasiveness and high cost [4]. However, with recent advances in imaging resolution and safety, MRA has emerged as a viable alternative to DSA in several clinical scenarios [5], owing to its greater accessibility and lower cost. In particular, time-of-flight (TOF) MRA is generally used, as it does not require contrast injection and highlights blood flow, underscoring its noninvasive nature [6]. This avoids both contrast-related allergic reactions and radiation exposure [7]. Consequently, TOF-MRA demonstrates relatively high sensitivity and is frequently used in clinical settings for cerebral aneurysm diagnosis and follow-up.

Despite advances in MRA, aneurysm identification remains challenging. Their asymptomatic nature necessitates careful and detailed examination, which significantly increases clinician workload. To alleviate this burden, deep learning-based methods have been developed to support clinical decision-making. Several studies have proposed U-Net-based frameworks that incorporate attention modules and residual connections in a patch-wise manner [8, 9]. However, these approaches often neglect clinical characteristics that may affect detection accuracy. Because aneurysms occupy a small fraction of brain volume, random patch-based strategies may degrade model performance by including irrelevant regions and introducing redundant information.

Other approaches aim to improve aneurysm detection by removing non-informative anatomical information. Joo et al. [10] and Noto et al. [11] proposed frameworks that leverage vessel segmentation and region of interest (ROI) extraction to focus on cerebral vasculature. Given that aneurysms occur exclusively along vessel walls, these methods suppress nonvascular brain regions and improve aneurysm detection. However, the high anatomical variability of cerebral vessels across individuals makes accurate vessel segmentation technically challenging and time-consuming. Moreover, obtaining and validating vessel annotations through expert confirmation requires substantial clinical effort.

This highlights a critical trade-off: context-agnostic random patch methods are inefficient, while annotation-heavy full segmentation methods are clinically impractical. To address this dilemma, we propose a novel framework to intelligently navigate the middle ground. The primary contributions include:

1. A dual-stream synergistic framework that separates aneurysm candidate detection from fine-grained morphological segmentation. This architecture allows each stream to be optimized, with their outputs conditionally fused to generate a robust final prediction [12].
2. Landmark-guided localization utilizing sparse anatomical priors, specifically the coordinates of 18 clinically relevant vessel locations, to guide detection. This strategy provides essential spatial context without the prohibitive annotation costs of full-vessel segmentation.
3. An adaptive loss function for landmark localization that adjusts dynamically based on aneurysm presence or absence, encouraging model learning of clinically meaningful spatial relationships between landmarks and lesions.
4. Evaluation of a large-scale clinical dataset demonstrates a competitive balance between detection sensitivity and false-positive (FP) rates, thereby offering a practical and effective solution for automated IA detection.

By explicitly focusing on clinically relevant vessel regions and avoiding full-vessel segmentation, our method balances annotation efficiency with high diagnostic performance, thus supporting practical clinical applicability.

Materials and Methods

The study was conducted in accordance with the principles of the Declaration of Helsinki. The study protocol was approved by the Institutional Review Board (IRB) of Severance Hospital (IRB Nos. 4–2022-0261, 4–2024-0336, and 4–2024-1015), which waived the requirement for informed consent owing to the retrospective study design.

Study Dataset and Preparation

Patient Cohort and Data Acquisition

The dataset was retrospectively collected from 880 patients who underwent three-dimensional (3D) TOF-MRA of the brain at Severance Hospital, Neurosurgery, between January 2006 and December 2023. Patients were excluded if they had (1) missing image slices, (2) non-saccular aneurysms, (3) ruptured aneurysms or hematomas that distorted the surrounding brain anatomy, and (4) other severe lesions affecting cerebral vascular morphology. After applying these criteria, 25 patients were excluded, yielding a primary cohort of 855 patients (79 aneurysm-negative and 776 aneurysm-positive).

To ensure robust evaluation of less frequent aneurysm locations (e.g., the posterior circulation) underrepresented in the initial test split, we added an additional 200 cases (20

aneurysm-negative and 180 aneurysm-positive) not used for model training. Therefore, the final dataset included 1055 cases, comprising 99 aneurysm-negative and 956 aneurysm-positive scans. The data inclusion and exclusion processes are summarized in Fig. 1.

Imaging was performed on scanners from three manufacturers (Philips, GE Medical Systems, and Siemens) across 15 different models, with field strengths of 1.5 T or 3 T. The acquisition parameters are summarized in Table 1.

Table 1 Imaging parameters of the TOF-MRA dataset

Parameter	3-T scanners	1.5-T scanners
TR	17.5–34.8 ms	19.3–30.0 ms
TE	2.3–4.0 ms	2.5–7.2 ms
Flip angle	14–20°	20–25°
Slice thickness	0.4–0.4 mm	0.6–1.4 mm

TOF-MRA time-of-flight-magnetic resonance angiography, *TR* repetition time, *TE* echo time

Ground Truth Annotation

Ground-truth labels were established using a rigorous consensus process. Three clinical experts (two neurosurgeons and one radiologist) independently reviewed each TOF-MRA scan to determine IA presence, number, and location. Disagreements between the two neurosurgeons were adjudicated by the radiologist.

To maximize the data utilization, a specialized annotation strategy was employed for patients with prior aneurysm treatment (e.g., coiling or clipping). The treated regions were annotated using a distinct label class (values in the 100 s). During training, patches containing these labels were excluded to prevent treatment-related artifacts from affecting model learning, while retaining usable portions of the scans [13].

In addition to aneurysm segmentation, experts annotated the center points of 18 major cerebral vessel regions where IAs frequently occur. These landmarks, listed in the Appendix, provided anatomical priors for the landmark-guided detection model. All manual segmentation and annotation were performed using ITK-SNAP (http://itksnap.org).

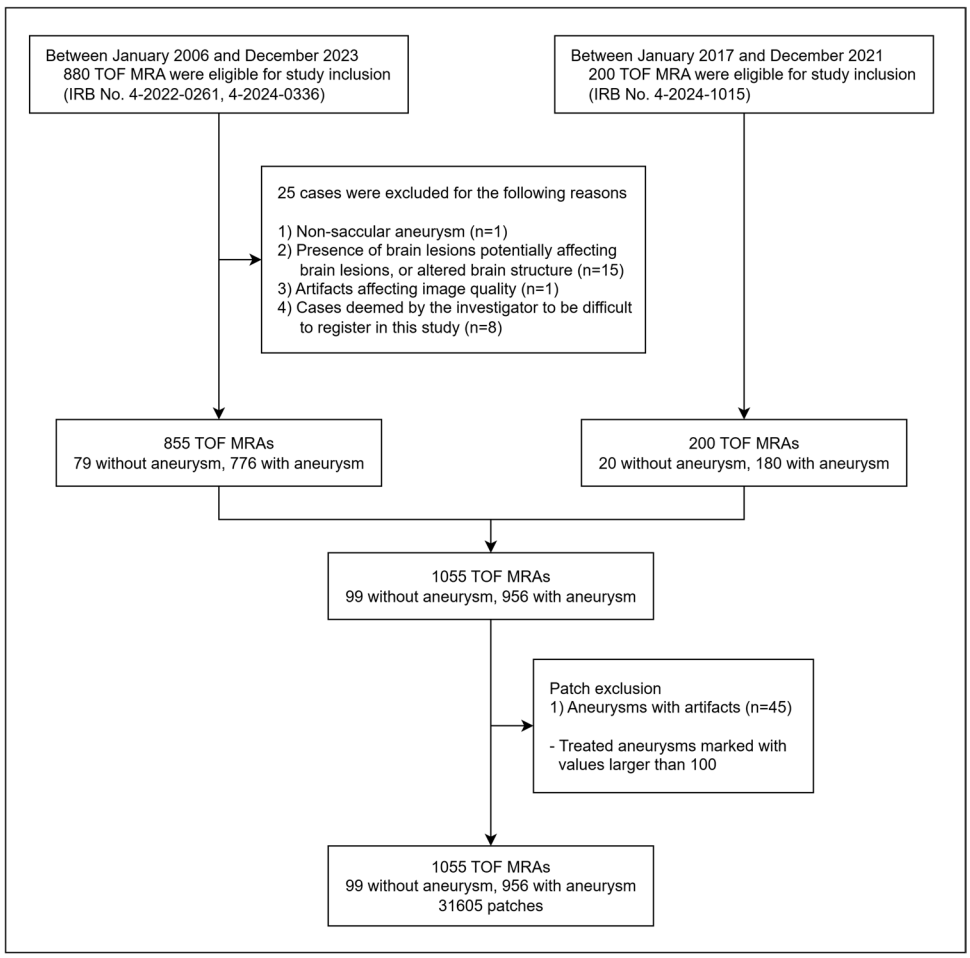

Fig. 1 Flowchart showing data inclusion and exclusion. TOF, time-of-flight; MRA, magnetic resonance angiography; IRB, institutional review board

Image Preprocessing Pipeline

All TOF-MRA scans underwent a standardized preprocessing pipeline before being input into the deep learning models, as follows:

1. Reorientation: Image volumes were reoriented to the right-anterior–superior coordinate system to ensure consistent anatomical alignment.
2. Histogram standardization: Applied to normalize intensity values across scans from different scanners and protocols to reduce inter-scan variability [14].
3. Intensity normalization: Voxel intensities were scaled to a range of [− 1, 1].
4. Resizing: Volumes were resized to (480, 480, and 160) voxels to streamline model training. By not resampling to a uniform voxel spacing, this approach intentionally preserves inherent variability in image resolution across the dataset, thus compelling the model to learn features robust to such differences, which are commonly seen in real-world clinical data.

Dual-Stream Synergistic Framework for Aneurysm Detection

Architectural Overview

The proposed dual-stream, coarse-to-fine framework synergistically combines targeted candidate detection with precise boundary segmentation (Fig. 2). The architecture consists of two parallel pathways that concurrently process input MRA data.

- Stream 1: Landmark-guided candidate detection: This stream operates in a coarse-to-fine manner to identify high-probability candidate aneurysm regions. Key vascular landmarks are first localized to guide a patch-based detection model. This stream primarily aims to maximize detection sensitivity.
- Stream 2: Fine-grained boundary delineation: This stream processes the full MRA volume to produce highly accurate segmentation masks and provides precise morphological details of potential lesions.

Finally, the outputs from these two streams are then combined using conditional fusion logic to synthesize a high-sensitivity candidate list from Stream 1 with high-precision boundaries from Stream 2 to generate the final output.

Stream 1: Landmark-Guided Candidate Detection

ROI Landmark Localization To efficiently guide the detection model to clinically relevant areas, a landmark localization network was developed to predict the 3D coordinates of 18 predefined vascular regions (Fig. 3). The network architecture is based on the U-Net framework, featuring a symmetric encoder-decoder structure with skip connections [15]. The encoder path consists of sequential blocks of convolution, instance normalization, dropout, and parametric ReLU (PReLU) activation, followed by max pooling. Rather than

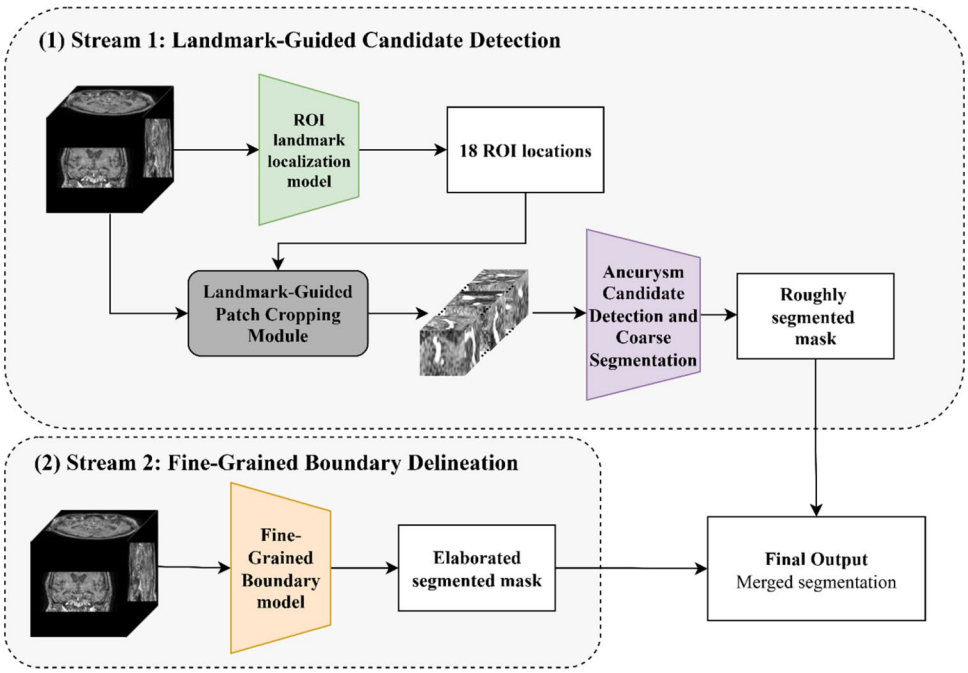

Fig. 2 Workflow overview of the proposed model. **1** Aneurysm detection consists of two staged models: ROI localization and detection. **2** The detailed segmentation model predicts a more elaborate mask. The final output from **1** and **2** is merged. ROI, region of interest

directly regressing coordinates, the model generates probabilistic heatmaps for each landmark to improve training stability and localization accuracy [16, 17].

Adaptive Loss for Landmark Localization We introduce a novel adaptive loss for ROI localization to encourage learning of meaningful spatial correspondences between aneurysms and vascular structures. The loss function is based on the mean squared error (MSE) between the predicted and ground-truth locations. During training, in the absence of an aneurysm, the model predicts the geometric center of each vessel region. For aneurysm-positive cases, the model identifies the center of the aneurysm and computes its Euclidean distance from the nearest vessel ROI. The MSE is used to minimize the distance when the aneurysm was associated with the vessel ROI and to penalize mismatches by increasing the distance otherwise, adding -MSE.

The loss L_{adaptive} is based on the MSE and is calculated as follows:

$$L_{\text{adaptive}} = \frac{1}{18}\sum_{i}^{18}\left(1-y_{\text{ia}}\right)\bullet MSE\left(P_{\text{i}},C_{\text{v,i}}\right) + y_{\text{ia}}\bullet\left(\lambda\bullet MSE\left(P_{\text{i}},C_{\text{a,i}}\right)+sim\left(P_{\text{i}},C_{\text{v,i}}\right)\right)$$

where i is the index for one of the 18 landmarks, P_{i} is the predicted coordinate, $C_{\text{v,i}}$ is the ground-truth center of the vessel region, and $C_{\text{a,i}}$ is the ground-truth center of the aneurysm. The binary indicator y_{ia} equals 1 if an aneurysm is present and associated with landmark i, and 0 otherwise. The term $sim(\bullet,\bullet)$ represents a similarity metric (e.g., a negative exponential of the distance) to penalize predictions far from the vessel center, even in positive cases. The weighting hyperparameter λ is empirically set to 2 to place greater emphasis on aligning the predictions with each other.

In practice, for aneurysm-negative cases ($y_{\text{ia}} = 0$), the loss minimizes the distance between the prediction and the vessel center. For aneurysm-positive cases ($y_{\text{ia}} = 1$), the loss primarily minimizes the distance to the aneurysm center while also encouraging proximity to the vessel center. These dual objectives guide the model to learn the typical spatial distribution of aneurysms around major vessels.

Landmark-Guided Patch Cropping Module Based on the coordinates predicted by the landmark localization model, a patch-cropping module is used to extract regions of interest (ROIs) for subsequent detection. For each predicted landmark, a $64\times64\times64$ voxel patch is cropped and centered on the predicted coordinates. This step effectively removes irrelevant extravascular tissue, allowing the detection model to focus on information-rich vascular structures.

Aneurysm Candidate Detection and Coarse Segmentation The core of the candidate detection stream is a hybrid architecture that combines an UNETR encoder with a feature-pyramid network (FPN) decoder [18] (Fig. 4). This design leverages the strengths of both transformers and convolutional neural networks (CNNs). The vision transformer backbone of UNETR [19] captures long-range global contextual features within each patch, while the FPN decoder handles multi-scale objects by generating predictions at different feature resolutions.

To compensate for the loss of global spatial context inherent in patch-based processing, 3D coordinate information is explicitly provided to the model. A coordinate map corresponding to each patch is concatenated channel-wise with the image data, making the model "coordinate-aware."

A two-stage refinement cycle is implemented to improve specificity and reduce FPs:

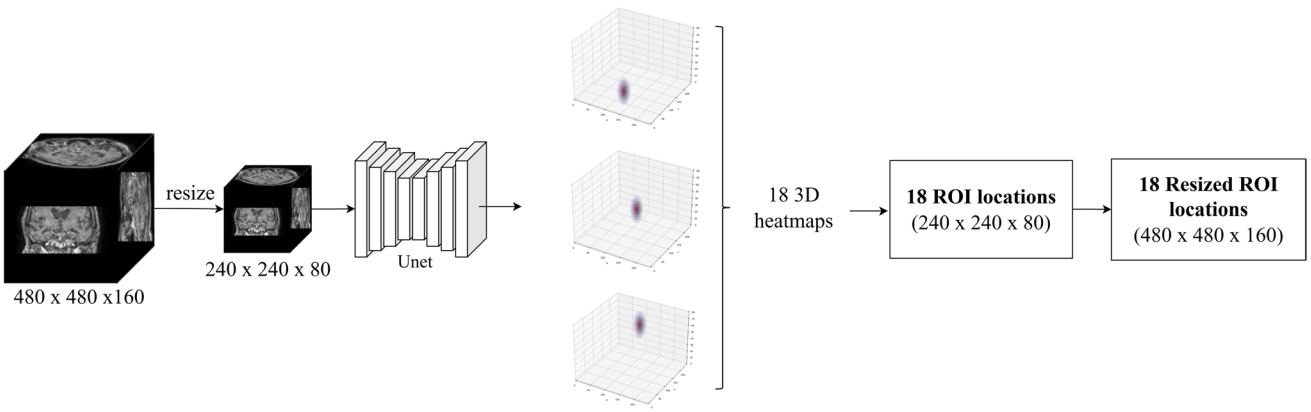

Fig. 3 Architecture of the proposed ROI landmark localization model. The model produces 18 heatmaps and generates ROI locations. These locations are fed back into the model for training; however, the final prediction restores the location to a size of $480\times480\times160$. 3D, three-dimensional; ROI, region of interest

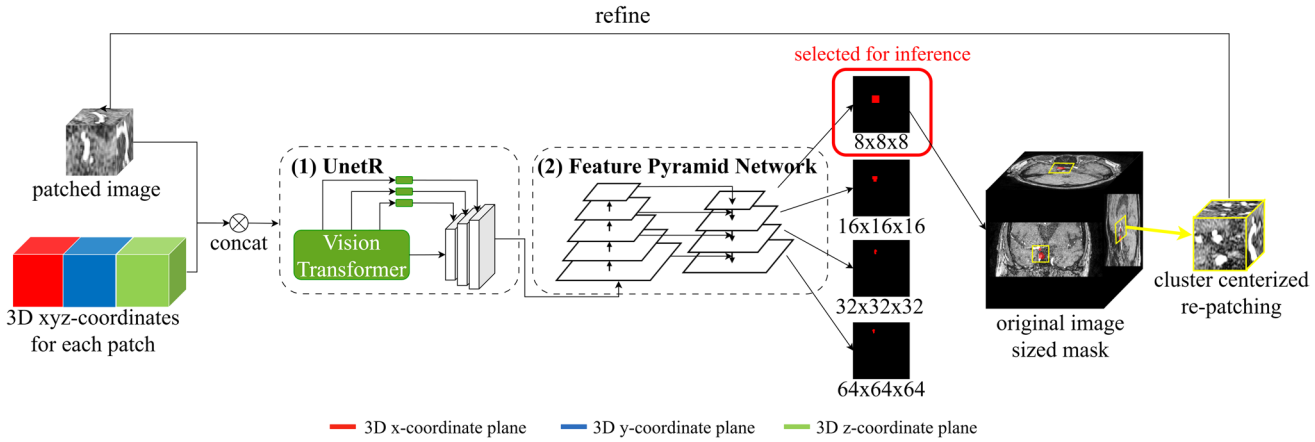

Fig. 4 Full detection model training pipeline using UnetR and feature pyramid network. The network refines the segmentation across multiple resolutions, while an 8 × 8 × 8-sized output is selected for inference and to reconstruct the full-sized mask. Each cluster in the mask is re-patched with cluster center alignment and refined by repeated training. 3D, three-dimensional

1. Initial prediction: The model first generates coarse segmentation predictions for landmark-guided patches.
2. Candidate refinement: Patch-level predictions are reassembled into a full-volume mask, and connected components (clusters) in this mask are identified as initial aneurysm candidates. Each candidate cluster is then used to define a newly centered patch. These candidate-centered patches are fed back into the same detection model for a second pass, which effectively acts as a classifier to prune FPs. The final output of this stream is a refined mask of high-confidence aneurysm candidates.

Stream 2: Fine-Grained Boundary Delineation

To achieve high-precision segmentation of the final aneurysm masks, a separate model is operated in parallel on the full, uncropped MRA volume. Specifically, we used an nnU-Net framework with a residual encoder (ResEnc) configuration [20, 21]. This architecture was chosen because nnU-Net is widely considered a state-of-the-art, self-configuring baseline for medical image segmentation, consistently delivering top-tier performance across diverse datasets and modalities without task-specific manual tuning [22]. This model receives the entire preprocessed image as input and is optimized to produce segmentations with precise and accurate boundaries.

Synergistic Output Fusion

The final prediction is generated by conditionally merging the outputs of the two parallel streams. This fusion logic establishes a clear hierarchy: the candidate detection stream (Stream 1) acts as the primary driver of sensitivity, while the boundary delineation stream (Stream 2) serves as a conditional refiner to enhance precision. The fusion rules are as follows:

1. Condition 1 (Overlap): If a candidate mask from Stream 1 overlaps with a segmentation mask from Stream 2, the final output for that region is obtained by element-wise multiplication of the two masks. This method uses the precise boundary from Stream 2 to refine the shape of the confirmed candidate from Stream 1.
2. Condition 2 (Detection only): If a candidate mask exists only in Stream 1 (no corresponding overlap in Stream 2), it is retained in the final output. This ensures that true aneurysms detected by the more sensitive candidate model are not discarded if the more conservative segmentation model fails to delineate them.
3. Condition 3 (Segmentation only): If a segmentation mask exists only in Stream 2 (without a corresponding candidate in Stream 1), it is discarded. This "gating" mechanism is critical, as it prevents potential FPs generated by the segmentation model.

Experimental Setup and Evaluation

Implementation Details

The dataset was randomly split into training and testing sets in an 80:20 ratio, yielding 695 training cases (29 aneurysm-negative, 666 aneurysm-positive) and 160 cases in the first internal test set (Test Set 1). The supplementary 200 cases formed the second internal test set (Test Set 2). Each model in the framework was trained separately using task-specific optimization parameters on NVIDIA RTX 4090 GPUs.

- ROI landmark localization: The model was trained using the AdamW optimizer with an initial learning rate of 1×10^{-4} and a dropout rate of 0.1. A step-decay learning rate scheduler was applied with a gamma of 0.8 per epoch.
- Candidate detection: Due to the high-class imbalance in patch-wise detection, we used a composite loss function of focal and Tversky losses to improve robustness [23]. The model was trained using the AdamW optimizer and a cosine annealing warm restart scheduler for 1000 epochs, with an initial learning rate of 1×10^{-3}.
- Boundary delineation (nnU-Net): This model was trained using a hybrid loss function combining Dice and cross-entropy losses. We used the Adam optimizer with an initial learning rate of 3×10^{-4}. The learning rate was reduced by a factor of 5 if the validation loss did not improve for 30 consecutive epochs.

Evaluation Protocol

We assessed the framework's performance using a comprehensive set of metrics.

- ROI localization sensitivity: To evaluate the effectiveness of the landmark localization model, we calculated the proportion of ground truth aneurysms fully contained within the $64 \times 64 \times 64$ patches generated from the predicted landmark coordinates.
- Lesion-wise sensitivity: This metric evaluated the model's ability to correctly identify aneurysms. It was calculated as the number of correctly detected aneurysms divided by the total number of ground truth aneurysms.
- FP rate: To measure the model's specificity and potential clinical burden, we calculated the average number of FP detections per patient scan.
- Dice similarity coefficient (DSC): The DSC was used to evaluate the spatial overlap and accuracy of the segmentation masks for correctly detecting lesions. The DSC was calculated for the output of the detection model alone, segmentation model alone, and final fused output.
- Hausdorff Distance 95% (HD-95): The HD-95 was used to assess the boundary errors of segmentation performance. Unlike DSC, HD-95 is sensitive to outliers; hence, it can be evaluated robustly. For further evaluation of the segmentation performance, HD-95 was calculated for final fused output only.
- True positive (TP) criterion: A predicted lesion was considered a TP if the Euclidean distance between the center of mass of the mask and the center of mass of the corresponding ground-truth mask was < 3 mm in the voxel space, which is a standard criterion used in similar studies [24, 25].

Table 2 Characteristics of the internal test sets

Characteristics	Internal Test Set 1	Internal Test Set 2
Number of scans (*N*)	**160**	**200**
Aneurysm-negative	50	20
Aneurysm-positive	110	180
Number of aneurysms (*N*)	**155**	**249**
Aneurysm size (mean ± std, mm)	**5.89 ± 2.41**	**6.29 ± 4.18**
≤ 3 mm (*N*)	6	15
3–5 mm (*N*)	59	94
5–7 mm (*N*)	53	84
> 7 mm (*N*)	37	56
Aneurysm locations		
ICA (*N*)	108	137
MCA (*N*)	30	55
ACA (*N*)	3	11
ACOM (*N*)	12	12
BA (*N*)	2	22
Posterior circulation (*N*)	0	12

std standard deviation, *ICA* internal carotid artery, *MCA* middle cerebral artery, *ACA* anterior cerebral artery, *ACOM* anterior communicating artery, *BA* basilar artery

Statistical Methods

All statistical analyses were performed using Python software (version 3.12.9). Associations between lesion-wise sensitivity and FP rate across models were explored using Spearman's rank correlation coefficient (ρ). Differences in lesion-wise sensitivity across vascular territories were assessed using chi-square and Fisher's exact tests, as appropriate. Statistical significance was defined as two-sided $\rho < 0.05$.

Results

Dataset Characteristics

The two internal test sets comprised a wide range of aneurysm sizes and locations (Table 2). Test Set 1 contained 160 scans of 155 aneurysms, and Test Set 2 included 200 scans of 249 aneurysms. The mean aneurysm size was approximately 6 mm in both sets, ranging from 2 to > 32 mm. Most aneurysms were located in the internal carotid artery and middle cerebral artery, consistent with their clinical prevalence. Test Set 2 was specifically curated to include more

aneurysms in the basilar artery and posterior circulation, which were underrepresented in Test Set 1.

Performance of the Proposed Framework

Table 3 summarizes the overall performance of the proposed dual-stream framework. The ROI landmark localization model demonstrated high effectiveness, with sensitivities of 0.97 and 0.98 in Test Sets 1 and 2, respectively, indicating the successful generation of patches encompassing nearly all ground-truth aneurysms.

The final fused output of the framework achieved a lesion-wise sensitivity of 0.87 with 1.23 FPs/case on Test Set 1, and a sensitivity of 0.82 with 1.17 FPs/case on Test Set 2. Regarding segmentation accuracy, the final merged output produced a dice similarity coefficient (DSC) of 0.68 in Test Set 1 and 0.70 in Test Set 2, demonstrating a substantial improvement over the coarse segmentation from the detection model alone (DSC of 0.28 and 0.34) and reflecting a successful fusion with high-quality masks from the fine-grained segmentation model (DSC values of 0.80 and 0.78).

Table 4 depicts the model's performance stratified by aneurysm size and location. The sensitivity was positively correlated with aneurysm size, with notably lower values for small aneurysms (≤ 3 mm) (0.67 and 0.40) compared with larger aneurysms (> 7 mm) (0.92 in Test Set 1). Conversely, detection was highly consistent across different anatomical locations.

Performance Comparison with State-of-the-Art Approaches

Benchmarking of the proposed model against those of prior studies (Fig. 5) revealed a general trade-off between sensitivity and false-positive rates in the literature. Many studies reporting sensitivity > 0.90 also exhibit high FP rates (e.g., > 2.5), while those with very low FP rates often achieve sensitivities < 0.75 [8, 9, 24, 26]. For instance, Joo et al. [10] achieved a high sensitivity of 0.92, with an exceptionally low FP rate of 0.123, representing state-of-the-art performance. In comparison, our model achieved a slightly lower sensitivity (0.87) with a higher FP rate (1.23), placing it in a competitive position that prioritizes high sensitivity while maintaining a clinically manageable FP burden.

Table 3 Overall detection and segmentation performance on internal test sets

Metric	Internal Test Set 1	Internal Test Set 2
ROI localization sensitivity	0.97 (151/155)	0.98 (240/246)
Lesion-wise sensitivity	0.87 (135/155)	0.82 (203/249)
FP rates	1.23 (196/160)	1.17 (234/200)
HD	1.57 mm	1.74 mm
DSC (final merged output)	0.68	0.70
DSC (detection model only)	0.28	0.34
DSC (segmentation model only)	0.80	0.78

ROI region of interest, *FP* false-positive, *DSC* dice similarity coefficient, *HD* Hausdorff distance.

Table 4 Lesion-wise sensitivity stratified by aneurysm size and location

Stratification	Category	Sensitivity (Test Set 1)	Sensitivity (Test Set 2)
By size	≤ 3 mm	0.67 (4/6)	0.40 (6/15)
	3–5 mm	0.85 (50/59)	0.81 (76/94)
	5–7 mm	0.89 (47/53)	0.90 (76/84)
	> 7 mm	0.92 (34/37)	0.80 (45/56)
By location	ICA	0.86 (93/108)	0.80 (109/137)
	MCA	0.93 (28/30)	0.84 (46/55)
	ACA	0.67 (2/3)	0.82 (9/11)
	ACOM	0.83 (10/12)	0.92 (11/12)
	BA	1.00 (2/2)	0.82 (18/22)
	Posterior circulation	N/A	0.83 (10/12)

ICA internal carotid artery, *MCA* middle cerebral artery, *ACA* anterior cerebral artery, *ACOM* anterior communicating artery, *BA* basilar artery

Ablation Study

To illustrate the improvement on the performance metrics from the proposed structure, we evaluated our framework by each stream on Internal Test Sets 1 as shown in Table 5. With the use of Stream 2 only, it achieved DSC of 0.80 and FP rate of 0.14. This is considered to be a very accurate performance, but the sensitivity has reached only 0.73. Stream 1 achieved the highest sensitivity of 0.87, but the DSC value produced 0.28. Therefore, fusion of the two streams adjusted DSC values, resulting in 0.87 sensitivity, 1.23 FP rate, and 0.68 DSC.

Discussion

In this study, we developed and validated a dual-stream deep learning framework for automated IA detection and segmentation using TOF-MRA data. By integrating a landmark-guided candidate detection stream with a parallel fine-grained segmentation stream, our method achieved a strong balance between lesion-wise sensitivity (0.82–0.87) and clinically acceptable FP rates (1.17–1.23 per case). The key contribution is its ability to leverage sparse anatomical priors to guide detection, thereby avoiding impractical

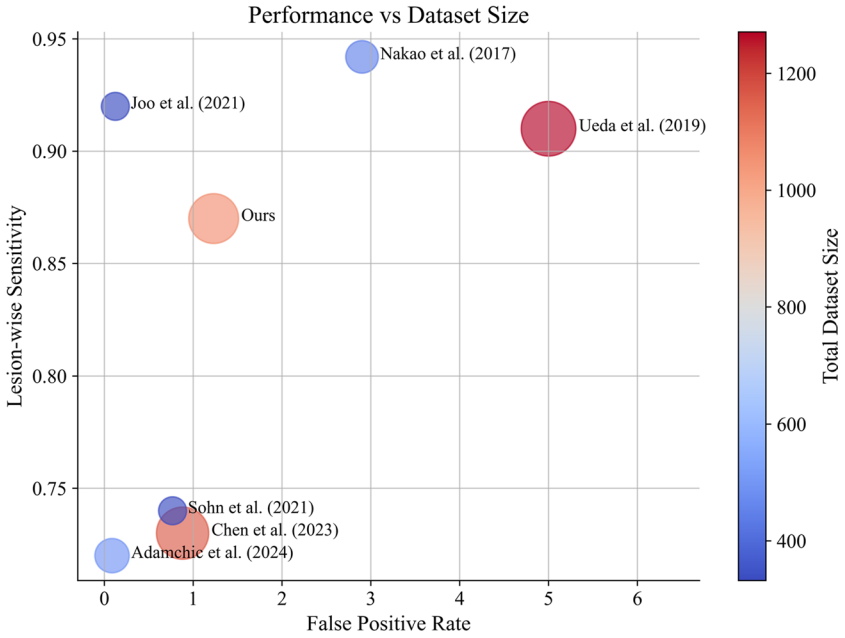

Fig. 5 Comparisons of overall lesion-wise sensitivity and false-positive rates between the model proposed in the present study and previous studies. A smaller false-positive rate and higher sensitivity are preferable

Table 5 An ablation study of proposed framework on each stream

Fused streams		Sensitivity	FP rate	DSC
Stream 1	Stream 2			
	✓	0.73	0.14	0.80
✓		0.87	1.23	0.28
✓	✓	0.87	1.23	0.68

FP false-positive, *DSC* dice similarity coefficient

full-vessel segmentation while providing sufficient context to outperform simple patch-wise methods.

Recent advances in AI have driven substantial progress across a wide range of medical fields. AI-based systems are increasingly used to support clinical workflows, assisting with tasks such as disease classification, lesion detection, segmentation, and even treatment planning. These advances reflect a broader shift in radiology toward AI-assisted interpretation, where AI systems increasingly function as complementary tools that enhance diagnostic accuracy and reduce oversight. This trend is also evident in neurovascular imaging, where the inherent complexity of the cerebral vasculature, the small size of many aneurysms, and substantial inter-reader variability make automated assistance highly valuable. Consequently, the integration of deep learning-based support systems into TOF-MRA workflows has gained growing clinical interest, reinforcing the relevance of developing robust IA detection models such as the framework proposed in this study.

The landmark localization model is the central component of the proposed framework. Its ability to correctly position patches over the ground truth aneurysms in > 97% of cases validated this guiding mechanism and represents a significant step toward clinical application. Although full-vessel segmentation methods have demonstrated excellent performance [26], the extensive annotation effort required limits their scalability. Our landmark-based strategy drastically reduces this burden by annotating only 18 points per scan, making the creation of large-scale training datasets far more feasible. Compared with other landmark-based methods, such as that described by Di Noto et al. [11], who reported a sensitivity of 0.80, our framework demonstrated a notable improvement, underscoring the effectiveness of our synergistic architecture and novel adaptive loss function.

The study further benefits from the patch-based training strategy, which resolves patient-level class imbalance. Our cohort includes a substantially higher number of aneurysm-positive patients compared to the negative (956 positive vs. 99 negative patients). Each patient scan is sampled with 18 anatomical locations and labeled positive if more than half of an aneurysm volume is included. This ensured adequate exposure to negative samples during optimization, enabling the model to learn reliable discriminative patterns despite the extreme imbalance of data distribution.

The performance of our model relative to the existing literature highlights the well-known trade-off between sensitivity and specificity. Spearman's rho ($\rho = 0.54$, $n = 5$) suggested a moderate positive trend between sensitivity and FP rates, although the small number of previous studies [27] limits the strength of this inference. Although some methods reported lower FP rates [9, 10, 26], the operating point of our model is clinically valuable for second-reader or computer-aided detection systems. In such workflows, maximizing sensitivity to ensure that no lesions are missed

is often prioritized, provided the FP rate remains low enough to avoid "alarm fatigue" for the reviewing clinician. An average of approximately one FP result per scan is a manageable workload in routine diagnostic settings. Accordingly, our model has the potential to reduce missed aneurysms by reliably flagging candidate lesions while simultaneously providing detailed segmentation masks that delineate aneurysm morphology. This dual output can support clinicians not only in detecting subtle aneurysms but also in assessing whether a lesion may warrant further intervention or surgical consideration. Such capabilities highlight the model's clinical utility and suggest its broader potential for integration into medical workflows.

The key strength of the proposed framework was its robust and consistent performance across diverse anatomical locations. The chi-square test showed no significant differences in lesion-wise sensitivity across the six major vessel territories ($p = 0.610$ for Test Set 1; $p = 0.924$ for Test Set 2). This location-invariant robustness is particularly important for clinical applicability, as it demonstrates reliability even in anatomically complex and high-risk regions such as the anterior communicating artery, basilar artery, and posterior circulation [2]. Previous studies reporting high-rupture-risk regions were often limited to a median of only 14 aneurysms, whereas our dataset comprised 74 aneurysms. This substantially larger sample size enhanced the reliability of our findings and provided a more robust foundation for clinical translation.

These advantages are expected to provide substantial benefits in real-world clinical practice. Notably, the diagnostic performance of the proposed model exceeds the previously reported performance of physicians in detecting intracranial aneurysms on TOF-MRA [29, 30]. These findings suggest that the proposed model may serve as a clinically assistive tool in the diagnosis of IAs. Importantly, the framework can be flexibly incorporated into different clinical settings. In healthcare screening centers or primary care environments where specialized neurovascular expertise may be limited, our framework could be utilized as a preliminary interpretation or screening tool to highlight suspicious regions and reduce the risk of missed IAs. In contrast, in tertiary medical centers with sufficient expert staffing, our framework may function as a supportive second reader or double-check tool, enhancing diagnostic confidence while maintaining a manageable false-positive burden. Hence, the integration of automated systems is designed to support clinicians, rather than replace clinical judgment.

However, our results revealed a clear performance dependency on aneurysm size. The sensitivity for aneurysms ≤ 3 mm was substantially lower than that for larger lesions. This is a common challenge in automated detection systems and reflects the inherent difficulty of identifying these tiny structures, which have a low signal-to-noise ratio and can easily be mistaken for vessel bifurcations or imaging artifacts.

This study has certain limitations. First, the validation was performed using data from a single institution. Although the dataset was large and included images from diverse scanners and protocols, external multicenter datasets are needed to establish the model's generalizability. Second, while the FP rate is acceptable for an assistive system, it is significantly higher than that of expert clinicians (approximately 0.03) [27]. Thus, the model is not yet suitable for fully autonomous diagnostic use and should be implemented as a decision-support tool within clinical workflows. In addition, we plan to further refine the model in future work to improve its performance for aneurysms smaller than 3 mm, aiming to minimize the performance gap across size categories. Last, even though the proposed model ensures balanced exposure to positive and negative samples, it is true that the original cohort does not fully reflect real-world prevalence where aneurysms are significantly less common.

Several future research directions are planned to address these limitations. First, we aim to investigate a dynamic patch-sizing mechanism to improve the detection of small aneurysms. This approach involves an initial coarse prediction to identify candidate regions, followed by the extraction of smaller, higher-resolution patches tailored to the lesion size, thereby improving feature representation of small aneurysms [1]. Additionally, as aneurysms ≤ 3 mm are rare [30], training data are inherently biased toward larger lesions. To mitigate this, we plan to use generative models such as generative adversarial networks (GANs) or diffusion models to synthesize anatomically plausible and realistic small aneurysms. Augmenting the training set with these synthetic examples is expected to improve the robustness and sensitivity of the model for this challenging but clinically critical subgroup. Finally, we intend to collect more clinically representative datasets with real prevalence distributions for validation, allowing more rigorous assessment of the model's clinical applicability.

Conclusions

This study introduced a dual-stream synergistic framework for the automated detection of IAs from TOF-MRA images. By combining landmark-guided ROI localization with parallel patch-wise detection and detailed segmentation, this method achieved competitive lesion-wise sensitivity with a low false-positive rate while avoiding the need for labor-intensive full-vessel segmentation. The framework demonstrated consistent, location-invariant performance across anatomically complex and high-risk vascular territories, highlighting its potential for broad clinical applicability. While the detection of very small aneurysms remains challenging, our promising approach provides a practical balance of detection accuracy, annotation efficiency, and clinical

utility. With further validation, this framework could be integrated into neurovascular imaging workflows as an assistive tool to support timely and accurate diagnosis of IAs.

Appendix

List of 18 major cerebral vessel landmarks

Left posterior inferior cerebellar artery (PICA), right PICA, left anterior inferior cerebellar artery (AICA), right AICA, left superior cerebellar artery (SCA), right SCA, basilar artery, left communicating internal carotid artery (ICA), right communicating ICA, left cavernous (CAV) ICA, right CAV ICA, left clinoid-ophthalmic ICA, right clinoid-ophthalmic ICA, left middle cerebral artery (MCA), right MCA, anterior communication artery, left distal anterior cerebral artery (ACA), right distal ACA.

Author Contribution DoyeonKim: conceptualization, methodology, software, validation, formal analysis, writing—original draft, and visualization. JieunPark: validation and writing—original draft. HyeonsikYang: conceptualization, methodology, and software. GiyounKim: methodology and software. JiyeonLee: writing—original draft. DonghyeonKim: supervision, project administration, and funding acquisition. Hyun Jin Han: resources, data curation, and supervision. KeunYoung Park: resources, data curation, and supervision. Minho Lee: writing—review and editing and supervision.

All authors have read and agreed to the published version of the manuscript.

Funding This research was financially supported by the Ministry of Trade, Industry and Energy, Korea, under the "Regional Innovation Cluster Development Program (R&D, P0025377)," supervised by the Korea Institute for Advancement of Technology (KIAT).

Data Availability The data and code in the current study is not applicable due to privacy, ethical, or legal issues.

Declarations

Ethics Approval This study was reviewed by the Institutional Review Board (IRB) of Severance Hospital and was granted an exemption from formal ethics approval as it used fully anonymized, retrospective data.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

Competing Interests The authors declare the following financial interests/personal relationships which may be considered potential competing interests: Author Donghyeon Kim has equity in Neurophet, Inc. Authors Doyeon Kim, Jieun Park, Hyeonsik Yang, Gi-youn Kim, Jiyeon Lee, Donghyeon Kim, and Minho Lee are employed by Neurophet, Inc. Hyun Jin Han and Keun Young Park participated as advisory professors in this study. No other author has reported a potential conflict of interest relevant to this article.

References

1. Jersey AM, Foster DM. Cerebral Aneurysm. [Updated 2023 Apr 3]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan-. Available at https://www.ncbi.nlm.nih.gov/books/NBK507902/. Accessed 29 August 2025.
2. Sanchez S, Miller JM, Samaniego EA: Clinical scales in aneurysm rupture prediction. Stroke: Vascular and Interventional Neurology. 4:Jan, 2024
3. Lee EJ, Lee HJ, Hyun MK, Choi JE, Kim JH, Lee NR, Hwang JS, Kwon JW: Rupture rate for patients with untreated unruptured intracranial aneurysms in South Korea during 2006-2009. J Neurosurg. 117:53-59, 2012
4. Howard BM, Hu R, Barrow JW, Barrow DL: Comprehensive review of imaging of intracranial aneurysms and angiographically negative subarachnoid hemorrhage. Neurosurg Focus 47:E20, 2019
5. de Boysson H, Boulouis G, Parienti JJ, Touzé E, Zuber M, Arquizan C, Dequatre N, Detante O, Bienvenu B, Aouba A, Guillevin L, Pagnoux C, Naggara O: Concordance of time-of-flight MRA and digital subtraction angiography in adult primary central nervous system vasculitis. AJNR Am J Neuroradiol 38:1917-1922, 2017
6. Boujan T, Neuberger U, Pfaff J, Nagel S, Herweh C, Bendszus M, Möhlenbruch MA: Value of contrast-enhanced MRA versus time-of-flight MRA in acute ischemic stroke MRI. AJNR Am J Neuroradiol 39:1710-1716, 2018
7. Andreucci M, Solomon R, Tasanarong A: Side effects of radiographic contrast media: Pathogenesis, risk factors, and prevention. Cairo: Hindawi Publishing Corporation, 2014. Available at https://doi.org/10.1155/2014/741018.
8. Ueda D, Yamamoto A, Nishimori M, Shimono T, shita S, Shimazaki A, Katayama Y, Fukumoto S, Choppin A, Shimahara Y, Miki Y: Deep learning for MR angiography: Automated detection of cerebral aneurysms. Radiology 290:187-194. 2019
9. Adamchic I, Kantelhardt SR, Wagner HJ, Burbelko M: Artificial intelligence can help detecting incidental intracranial aneurysm on routine brain MRI using TOF MRA data sets and improve the time required for analysis of these images. Neuroradiology 66:2195-2204, 2024
10. Joo B, Choi HS, Ahn SS, Cha J, Won SY, Sohn B, Kim H, Han K, Kim HP, Choi JM, Lee SM, Kim TG, Lee SK: A deep learning model with high standalone performance for diagnosis of unruptured intracranial aneurysm. Yonsei Med J 62:1052-1061, 2021
11. Di Noto T, Marie G, Tourbier S, Alemán-Gómez Y, Esteban O, Saliou G, Cuadra MB, Hagmann P, Richiardi J: Towards automated brain aneurysm detection in TOF-MRA: Open data, weak labels, and anatomical knowledge. Neuroinformatics 21:21-34, 2023
12. Wu L, Chen A, Salama P, Dunn KW, Delp J: An ensemble learning and slice fusion strategy for three-dimensional nuclei instance

segmentation. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 1883–1893, 2022
13. Lee DN, Li Y, Olsson LT, Hamilton AM, Calhoun BC, Hoadley KA, Marron JS, Troester MA. Image analysis-based identification of high risk ER-positive, HER2-negative breast cancers. Breast Cancer Res 26:177, 2024
14. Nyúl LG, Udupa JK, Zhang X: New variants of a method of MRI scale standardization," IEEE Trans Med Imaging19:143-150, 2000
15. Ronneberger O, Fischer P, Brox T: U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N, Hornegger J, Wells W, Frangi A (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Springer, Cham, 2015
16. Li J, Wang Y, Mao J, Li G, Ma R: End-to-end coordinate regression model with attention-guided mechanism for landmark localization in 3D medical images. In: Liu M, Yan P, Lian C, Cao X (eds) Machine Learning in Medical Imaging. MLMI 2020. Lecture Notes in Computer Science, vol 12436. Springer, Cham, 2020, https://doi.org/10.1007/978-3-030-59861-7_63
17. Kerfoot E, Clough J, Oksuz I, Lee J, King AP, Schnabel JA: Left-ventricle quantification using residual U-Net. In: Statistical Atlases and Computational Models of the Heart Atrial Segmentation and LV Quantification Challenges, 2018, vol 11395. Springer, Cham. https://doi.org/10.1007/978-3-030-12029-0_40
18. Lin T-Y, Dollár P, Girshick R, He K, Hariharan B, Belongie S: Feature pyramid networks for object detection. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 936–944, https://doi.org/10.1109/CVPR.2017.106
19. Hatamizadeh A, Tang Y, Nath V, Yang D, Myronenko A, Landman B, Roth HR, Xu D: UNETR: transformers for 3D medical image segmentation. 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2022. pp. 1748–1758, https://doi.org/10.1109/WACV51458.2022.00181
20. Isensee F, Petersen J, Klein A, Zimmerer D, Jaeger PF, Kohl S, Wasserthal J, Koehler G, Norajitra T, Wirkert S, Maier-Hein KH: nnU-Net: self-adapting framework for u-net-based medical image segmentation. In: Handels H, Deserno T, Maier A, Maier-Hein K, Palm C, Tolxdorff T. (eds) Bildverarbeitung für die Medizin 2019. Informatik aktuell. Springer Vieweg, Wiesbaden, 2019. https://doi.org/10.1007/978-3-658-25326-4_7
21. Isensee F, Wald T, Ulrich C, Baumgartner M, Roy S, Maier-Hein K, Jäger PF: "nnU-Net revisited: a call for rigorous validation in 3D medical image segmentation. Medical Image Computing and Computer Assisted Intervention – MICCAI 2024. MICCAI 2024. Lecture Notes in Computer Science, vol 15009. Springer, Cham, 2024. https://doi.org/10.1007/978-3-031-72114-4_47
22. Indrakanti AK, Wasserthal J, Segeroth M, Yang S, Nicoll AP, Schulze-Zachau V, Lieb J., Cyrlac M, Psychoglos M, Mutke MA: Multi-centric AI Model for Unruptured Intracranial Aneurysm Detection and Volumnetric Segmentation in 3D TOF-MRI. J Digit Imaging, 2025. https://doi.org/10.1007/s10278-025-01533-3
23. Du Nguyen Q and Thai HT, Crack segmentation of imbalanced data: The role of loss functions. Eng Struct 297: Dec, 2023. https://doi.org/10.1016/j.engstruct.2023.116988
24. Nakao T, Hanaoka S, Nomura Y, Sato I, Nemoto M, Miki S, Maeda E, Yoshikawa T, Hayashi N, Abe O: Deep neural network-based computer-assisted detection of cerebral aneurysms in MR angiography. J Magn Reson Imaging 47:948-953, 2018
25. Yang X, Blezek DJ, Cheng LT, Ryan WJ, Kallmes DF, Erickson BJ: Computer-aided detection of intracranial aneurysms in MR angiography. J Digit Imaging 24:86-95, 2011
26. Chen G, Yifang B, Jiajun Z, Dongdong W, Zhiyong Z, Ruoyu D, Bin D, Sirong P, Daoying G, Meng C, Yakang D, Yuxin L: Automated unruptured cerebral aneurysms detection in TOF MR angiography images using dual-channel SE-3D UNet: a multi-center research," Eur Radiol 33:3532-3543, 2023
27. Sohn B, Park KY, Choi J, Koo JH, Han K, Joo B, Won SY, Cha J, Choi HS, Lee SK: Deep learning-based software improves clinicians' detection sensitivity of aneurysms on brain TOF-MRA. AJNR Am J Neuroradiol 42:1769-1775, 2021
28. Chandra RV, Maingard J, Slater L-A, Cheung NK, Lai LT, Gall SL, Thrift AG and Phan TG, A Meta-Analysis of Rupture Risk for Intracranial Aneurysms 10mm or Less in Size Selected for Conservative Management Without Repair. Front Neurol 12:743023, 2022. https://doi.org/10.3389/fneur.2021.743023
29. Okahara M MD, Kiyosue H MD, Yamashita M MD, Nagatomi H MD, Hata H MD, Saginoya T MD, Sagara Y MD, and Mori H MD, Diagnostic Accuracy of Magnetic Resonance Angiography for Cerebral Aneurysms in Correlation With 3D-Digital Subtraction Angiographic Images: A Study of 133 Aneurysms. Stroke 33(7):1803-1808, 2002
30. Miki S, Hayashi N, Masutani Y, Nomura Y, Yoshiawa T, Hanaoka S, Nemoto M, Ohotomo K, Computer-Assisted Detection of Cerebral Aneurysms in MR Angiography in a Routine Image-Reading Environment: Effects on Diagnosis by Radiologists. AJNR Am J Neuroradiol 37(6):1038-43, 2016