This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2026.3735398

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2024.0429000

Precise Center Point Determination of Brain Vessel Branch Points and Endpoints After Deep Learning Object Detection

SAMET KAYA[1], BERNA KIRAZ[2]
[1]Department of Computer Engineering, Fatih Sultan Mehmet Vakif University, Istanbul, 34445, Turkey (e-mail: skaya@fsm.edu.tr)
[2]Department of Artificial Intelligence and Data Engineering, Istanbul Technical University, Istanbul, 34467, Turkey

Corresponding author: Samet Kaya (e-mail: skaya@fsm.edu.tr).

This work was supported by the Fatih Sultan Mehmet Vakif University Scientific Research Projects (SRP) Unit under grant number 22022B1Ç01D.

ABSTRACT Accurate localization of vascular landmarks, including branching points and endpoint(s), is fundamental for quantitative morphological analysis in neurovascular research and plays a critical role in the early detection of cerebrovascular pathologies, risk assessment, and treatment planning. Although deep learning-based object detectors have achieved high accuracy in identifying these landmarks, their bounding box outputs inherently lack the coordinate precision required for detailed morphometric studies such as three-dimensional vessel reconstruction and hemodynamic modeling. This study aims to address this limitation by proposing a novel hybrid methodology that integrates the robust detection capabilities of a deep learning model with the pixel-level precision of a targeted morphological post-processing stage. In the proposed approach, a Faster R-CNN model first identifies candidate regions by generating bounding boxes, and subsequently, a hit-or-miss transform (HMT) is applied within each bounding box to determine the precise center coordinates. The methodology was validated on 96 skeletonized slices from a single mouse brain—the High-Fat/High-Sucrose 13 (HFHS13) vasculature dataset, which exhibits significant topological heterogeneity—using a held-out test set (76 training / 20 test slices) with bootstrap 95% confidence intervals; because the slices are consecutive z-sections of one brain, we additionally verified performance under a leakage-free contiguous block split. We emphasize that the evaluation scope is limited to a single subject and that the method depends on high-quality upstream vessel segmentation and skeletonization. Experimental results demonstrate strong performance on the held-out test set, with precision rates of 99.86% for endpoint(s) and 99.82% for branching points, corresponding recall rates of 99.15% and 94.13%, and F1-scores of 99.50% and 96.89%, respectively (mean over three independent training runs). For the subset of correctly matched detections (true positives, determined via a distance-based matching criterion with a 5-pixel tolerance), the localization error was pixel-level, with a mean below 1 px, effectively converting coarse bounding boxes into precise pixel coordinates. Furthermore, our analysis revealed that the deep learning model demonstrates context-aware behavior by intelligently disregarding imaging artifacts, such as vessel truncations at image borders, which traditional morphological methods often misclassify as anatomical endpoint(s). This study presents a promising framework for extracting topological features from vascular networks, combining the contextual awareness of deep learning with the precision of mathematical morphology, though broader validation across datasets, species, and imaging modalities remains necessary to establish generalizability.

INDEX TERMS Deep learning, object detection, vessel analysis, center point localization, image processing, morphological analysis, hit-or-miss transform.

I. INTRODUCTION

SIGNIFICANT advances in medical imaging technologies in recent years have enabled the development of new methodologies for analyzing neurovascular structures. Accurate identification of morphological features within brain vascular networks, especially branching points and terminal endpoint(s), is crucial for understanding the pathophysiology of neurological diseases and guiding clinical diagnosis and

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2026.3735398

treatment [1]. Quantitative evaluation of these anatomical structures plays a decisive role in fundamental clinical processes such as the early detection of vascular pathologies, the optimization of treatment strategies, and the objective monitoring of therapeutic responses. Three-dimensional quantitative analysis of the vascular network is gaining increasing importance in the diagnosis and follow-up of cerebrovascular diseases (e.g., stroke, aneurysms, and vascular malformations) and neurodegenerative disorders [2], [3].

Traditional image processing techniques, such as edge detection, thresholding, morphological operators, and skeletonization, have long been used for vessel segmentation and morphometric analysis [4]. However, these approaches often require manual parameter optimization and struggle to produce consistent results, particularly with heterogeneous datasets [5]. Factors such as variability in image quality, low signal-to-noise ratio, contrast differences, and subject-to-subject anatomical variations significantly limit the clinical applicability of these methods. Furthermore, the inadequacy of traditional algorithms in terms of processing speed and scalability in large data sets, low levels of automation, and operator dependency are significant factors that prevent their widespread use in clinical routines [6]. Deep learning-based approaches, particularly convolutional neural networks (CNNs), have recently offered promising solutions to overcome these limitations [7]. When trained on large-scale and diverse medical image datasets, these models can automatically learn the complex features of vessel morphology, minimizing the need for human intervention. Compared to traditional methods, deep learning algorithms offer higher accuracy, sensitivity, and specificity in detecting vessel branching points and terminal endpoint(s), while demonstrating generalizable performance across different imaging modalities (e.g., Magnetic Resonance Angiography (MRA) and Computed Tomography Angiography (CTA), DSA) [8]. The high processing speed and automation capacity of these methods allow the rapid and standardized analysis of large patient cohorts [9].

The integration of deep learning-based vessel analysis systems into clinical applications has created new opportunities for neurovascular research and patient care. Automated vessel mapping and quantification systems help researchers better understand the natural progression of vascular pathologies and provide clinicians with objective and reproducible data for diagnostic decisions [10]. Moreover, the high-volume, standardized data generated by these technologies form a solid foundation for developing AI-assisted clinical decision support systems. Consequently, deep learning-based vessel analysis methods stand out as innovative tools that are shaping the future of neurovascular imaging and have the potential to create a paradigm shift in clinical neurology and neuroradiology [11].

Accurate and reliable detection of branching points and terminal endpoint(s) of vascular networks plays a central role in both basic scientific research and clinical applications. Branching points are essential for determining the topology of vascular networks and are key regions that directly influence blood flow distribution, hemodynamic stress, and vessel wall biology. It is well established that cerebral aneurysms and arteriovenous malformations frequently develop at these sites. Therefore, the precise mapping of branching points is critically important for the early detection of pathological processes and risk assessment [12]. Similarly, the terminal endpoint(s) of the vascular network represent the most distal regions where the oxygen and nutrient demands of tissues are met. The number, distribution, and morphological characteristics of these endpoint(s) serve as important biomarkers, particularly in the evaluation of microvascular pathologies such as small vessel disease and ischemic lesions [13].

In our previous studies [14], deep learning-based object detection models such as Faster R-CNN, Fast R-CNN, and RetinaNet were successfully applied to detect branching and endpoint(s) in skeletonized brain vessel images, achieving accuracy rates exceeding 90%. Subsequently, hyperparameter optimization was conducted to further enhance the performance of these models.

A systematic hyperparameter search, particularly on the Faster R-CNN model, resulted in optimization of parameters such as base learning rate, maximum number of iterations, and batch size, increasing the accuracy rate to over 98% [15]. This improvement demonstrates that, with appropriate tuning, deep learning-based object detection methods can yield highly reliable and generalizable results. While these object detection models performed their task with high precision by marking critical points with bounding boxes, these boxes themselves only provide general location information and do not provide the exact center coordinates required for detailed quantitative analysis of vessel morphology.

To provide a structured overview of the existing methodologies in vascular analysis and landmark detection, Table 1 summarizes the key studies, highlighting their contributions and limitations. This comparison clearly illustrates the gap that our proposed hybrid methodology aims to address: achieving pixel-level localization accuracy, with mean localization error below one pixel, while maintaining the robustness of deep learning-based detection.

In this study, we employed Faster R-CNN rather than single-stage detectors such as the YOLO family (including the recent YOLOv11). This choice is motivated by several technical considerations specific to our task. First, the vascular landmarks we aim to detect—branching points and endpoint(s)—occupy extremely small regions in the image, often only a few pixels in diameter. Two-stage detectors like Faster R-CNN are known to outperform single-stage architectures on small object detection tasks due to their region proposal mechanism, which allows for more refined feature extraction at multiple scales [16], [17]. Second, our application prioritizes localization precision over inference speed; the subsequent HMT refinement step requires accurate initial bounding boxes, and the higher localization accuracy of Faster R-CNN provides a more reliable foundation for this post-processing stage. While YOLO variants offer significant

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2026.3735398

speed advantages, this benefit is less critical in our offline analysis pipeline, where pixel-level localization accuracy is paramount. Furthermore, our previous comparative experiments [14] demonstrated that Faster R-CNN achieved superior detection performance on this specific dataset compared to other architectures, including RetinaNet.

Most studies in the literature either struggle with the noise sensitivity of purely classical methods or incur the high computational burden associated with end-to-end deep learning models (such as regression or heatmaps). It is important to position our approach relative to alternative paradigms for landmark localization, which are also summarized in Table 1 for completeness. Heatmap-based keypoint regression methods predict a probability distribution over the image and identify landmarks at local maxima; while effective for anatomical keypoints with consistent appearance, they require substantial training data and may struggle with the variable morphology of vascular junctions. Dedicated keypoint detection heads, such as those in CenterNet or CornerNet architectures, offer end-to-end learning but similarly demand large annotated datasets that are scarce in the neurovascular domain. Skeleton-graph methods extract landmarks through rule-based analysis of topological structures, achieving deterministic results but lacking robustness to segmentation noise. Our hybrid approach occupies a middle ground: it leverages the contextual learning capacity of object detection to identify candidate regions while delegating precise localization to a deterministic morphological operator, thereby combining learned robustness with geometric precision. This design choice reflects a pragmatic trade-off given the limited availability of pixel-level annotated vascular landmark datasets. This study presents a computationally efficient and pixel-precise solution by combining the detection power of deep learning with the precision of mathematical morphology.

In this context, the present study proposes a post-processing methodology to determine the precise center coordinates of branching points and terminal endpoint(s) within the bounding boxes identified by the deep learning model. The proposed method utilizes morphological image processing techniques to achieve pixel-level accuracy (mean <1 px) in center-point localization.

The accuracy of the detected centers was evaluated through a pixel-wise comparison with semi-automatically generated ground truth that was first obtained using traditional image processing techniques and subsequently refined by expert inspection, as described in [14]

The main contributions of this study can be summarized as follows:

- A novel hybrid methodology that integrates the robust object detection capabilities of Faster R-CNN with pixel-level precision of morphological Hit-or-Miss Transform (HMT) is proposed, which addresses the inherent localization inaccuracy in bounded boxes.
- In contrast to standard DL approaches that provide coarse regions, the presented methodology achieves pixel-level localization accuracy with a mean localization error below one pixel (<1px) for both branching points and endpoint(s), enabling high precision quantitative morphometric analysis.
- The proposed approach demonstrates superior robustness against imaging artifacts, specifically by correctly identifying and disregarding vessel truncations at image borders, which are often misclassified by traditional morphological methods.
- Through the work we summarize, we present a fully automated, end-to-end pipeline, which is validated through a case study that includes a complex mouse brain vasculature dataset, providing its potential for high-throughput neurovascular research without the need for manual intervention.

By integrating deep learning with traditional image processing, this approach aims to enhance both the accuracy and reliability of structural analysis of brain vessels, providing a solid foundation for 3D vessel reconstruction and advanced biomedical imaging analyses.

II. PROBLEM DEFINITION

Deep learning-based object detection algorithms have proven effective in identifying branching points and terminal endpoint(s) in cerebral vessels. Each detected point is indicated by a bounding box, which approximately defines the region where the corresponding feature is located. These bounding boxes not only illustrate the model's detection performance, but also facilitate the automatic localization of critical points. Figure 1 presents the branch and endpoint(s) detected by our deep learning-based method. As seen in the figure, bounding boxes only cover a general region and do not directly provide the exact center coordinates of the detected points. Additionally, although the detected boxes are small and tightly sized (their dimensions are quantified in Section IV), a box centroid does not necessarily coincide with the exact topological skeleton pixel, and in dense regions the boxes of neighbouring landmarks overlap. This complicates the determination of a precise center, particularly in regions with complex vascular structures or where detected points are located in close proximity.

To formalize the problem, we define the key terms used throughout this study. A *center point* refers to the single skeleton pixel that represents a topological landmark—specifically, the unique pixel identified by the HMT as satisfying the connectivity pattern of either an endpoint (exactly one foreground neighbor) or a branch point (three or more foreground neighbors). The output of our method is a set of predicted points $P = \{p_1, p_2, ..., p_n\}$ where each $p_i = (x_i, y_i, c_i)$ represents integer pixel coordinates and class label. Evaluation is framed as a bipartite matching problem between predicted points P and ground truth points G. We employ a greedy nearest-neighbor matching with a distance threshold of 5 pixels: each prediction is matched to its closest unmatched ground truth point of the same class if within the threshold; unmatched predictions are false positives, and unmatched ground truth points are false negatives. This protocol

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2026.3735398

does not explicitly handle duplicates beyond the one-to-one matching constraint—if multiple predictions fall near a single ground truth, only the closest is matched and others become false positives. We acknowledge that this matching protocol, while straightforward, may influence reported metrics; alternative protocols such as Hungarian matching could in principle yield different results. To quantify this, we note that the point annotations are 5×5 pixel boxes centered on each landmark, so the 5-pixel tolerance corresponds to approximately one box width on the 520×520 skeleton—a deliberately small neighborhood. We further performed a sensitivity analysis, sweeping the distance threshold over $\{3, 4, 5, 6, 8, 10\}$ pixels and comparing greedy matching against optimal (Hungarian) bipartite assignment. As reported in Section V (Table 4), the aggregate precision, recall, and F1 are invariant across this entire range and are identical for greedy and Hungarian matching, confirming that the reported figures reflect the detections themselves rather than the choice of matching protocol.

Another important issue is the potential overlap of the bounding boxes. Even after applying Non-Maximum Suppression (NMS) with a user-defined threshold, overlaps may persist in dense regions. This is particularly problematic at branching points and endpoint(s), where multiple bounding boxes may cover the same feature, reducing localization accuracy. Therefore, a robust method is required to automatically determine the exact centers of the branching points and endpoint(s) from the bounding boxes. The proposed approach must account for uncertainties such as variability in box sizes and overlaps to enable more accurate analysis of vascular topology and achieve more reliable results in medical applications.

III. METHODOLOGY

In this study, we develop a hybrid approach to precisely determine brain vessel branching points and terminal endpoint(s). This methodology combines the automation and generalization capabilities of deep learning with the high precision of traditional image processing techniques.

The foundation of this work is a high-quality ground truth dataset, the creation of which constituted a critical preliminary step, as detailed in our previous work [14]. To construct this dataset, we systematically applied traditional image processing techniques to the entire set of skeletonized images. Specifically, we employed the HMT using a comprehensive set of spatial filter kernels designed to identify all possible configurations of branching points and endpoint(s). The coordinates of every point detected by this automated process were then compiled. The initial ground truth points generated by morphological operations were independently verified by two domain experts. Any discrepancies between the experts or clear misidentifications (e.g., vessel border truncations labeled as endpoint(s)) were discussed and manually corrected to create a final, high-fidelity ground truth dataset. This rigorous verification process ensured that the final dataset was artifact-free and suitable to serve as a reliable *gold standard* for both training the deep learning model and evaluating the final performance of our hybrid methodology. The validated points were stored in COCO format to ensure compatibility with deep learning frameworks.

It should be noted that we acknowledge a methodological consideration regarding the use of HMT in both ground truth generation and inference, which we address in detail at the end of this section to maintain the flow of the methodological description.

In the second stage, a pre-trained Faster R-CNN model was employed to automatically detect branching points and endpoint(s) in new images. The Faster R-CNN architecture utilized a ResNet-50 backbone with Feature Pyramid Network (FPN), initialized with ImageNet-pretrained weights. The 96 skeletonized slices were partitioned into a training set of 76 slices and a held-out test set of 20 slices, and all reported metrics are computed on the held-out test set. Training used the standard detection pipeline with multi-scale resizing (shortest edge sampled from 640–800 px, maximum 1333 px) and random horizontal flipping. The model was trained for 14,000 iterations with a batch size of 2, using SGD with momentum 0.9, weight decay 0.0001, and a base learning rate of 0.001 held constant after a 1,000-iteration linear warm-up (no step decay); gradient-norm clipping at 1.0 was applied for stability given the very small landmark boxes. For bounding-box generation around point annotations, we used fixed-size boxes of 5×5 pixels centered on each ground-truth point, as vascular landmarks occupy only a few pixels on the skeleton. At inference, we applied a score threshold of 0.7 to filter low-confidence detections. The model's raw outputs were processed using Non-Maximum Suppression (NMS) with an Intersection over Union (IoU) threshold of 0.7, matching the value fixed in the released pipeline configuration. This process resulted in a final set of bounding boxes. However, these boxes provide coarse localization of the regions of interest. To overcome this limitation, the traditional spatial filtering methods used for ground truth generation were reapplied within each bounding box as a post-processing step. This enabled the automatic and high-precision determination of the exact center coordinates of the branches and endpoint(s) within the confined areas of the bounding boxes.

In the final stage, the points identified by our hybrid approach were compared pixel-wise with the ground truth data. This comparison enabled a comprehensive assessment of the method's performance using both qualitative observations and quantitative metrics.

The overall workflow is illustrated in Figure 2. Initially, the trained deep learning model detects candidate points on the skeletonized image by generating bounding boxes. Subsequently, a post-processing step based on the HMT is applied within each bounding box to localize the center points precisely. Finally, the resulting coordinates are compared with the ground-truth annotations to evaluate the performance of the proposed method.

All experiments were conducted on a system with an NVIDIA GeForce RTX 3080 GPU, 64 GB of RAM, and an

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2026.3735398

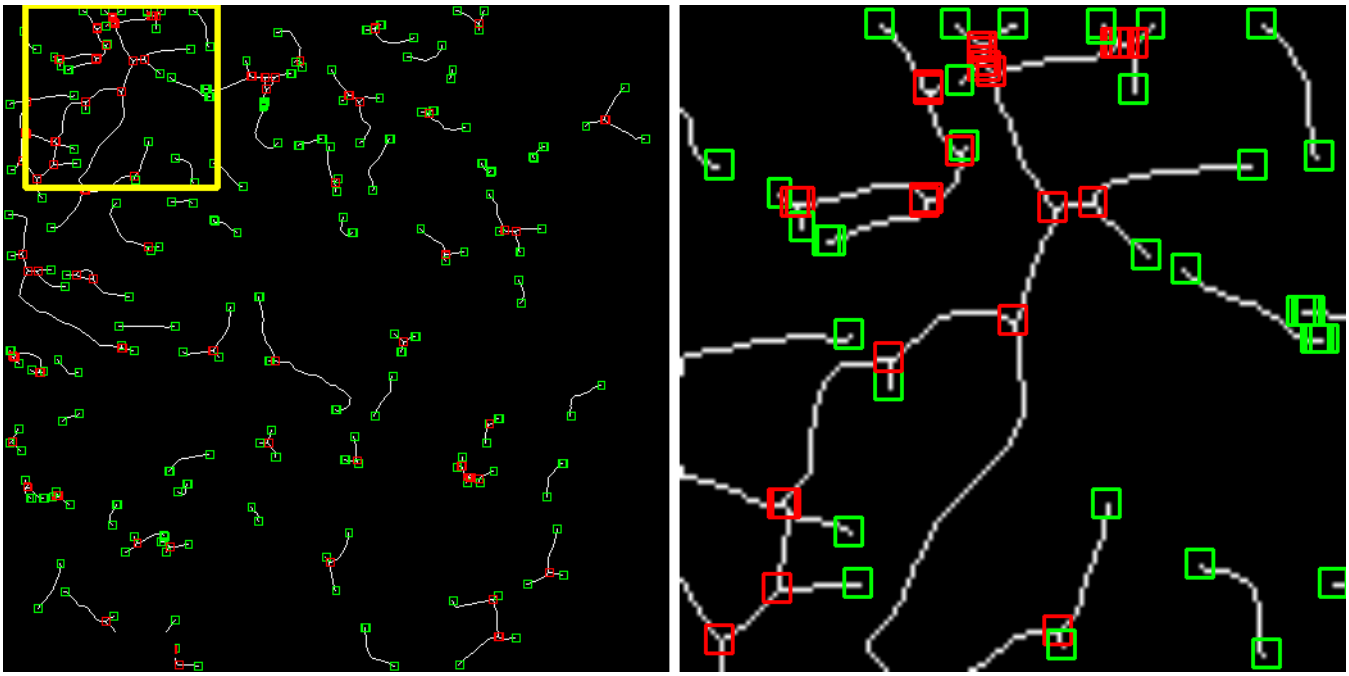


FIGURE 1. **Deep-learning detector output on a held-out slice, illustrating why a center-point refinement stage is needed. Left: all detected bounding boxes (endpoints green, branching points red) over the full 520×520-pixel skeletonized slice; the yellow rectangle marks the 150×150-pixel region enlarged at right. Right: in this dense region several 5×5 detection boxes for neighbouring landmarks overlap, and a box centroid need not fall on the exact skeleton pixel—the ambiguity that the hit-or-miss refinement resolves.**

Intel Core i7-10700K CPU. The average inference time for a single 520×520 pixel image, including both the Faster R-CNN detection and the HMT refinement, was approximately 150 milliseconds. Breaking down the computational cost: the Faster R-CNN detection stage accounts for approximately 135 ms (90% of total time), while the HMT refinement across all detected bounding boxes requires approximately 15 ms (10% of total time). The standard deviation across images was ±12 ms, with variation primarily dependent on the number of detected landmarks per image (range: 15–341 landmarks). This demonstrates the potential for practical application in high-throughput analysis.

Methodological Note on Ground Truth Generation: As mentioned earlier, we acknowledge a methodological consideration regarding the use of HMT in both ground truth generation and inference. While this creates a potential circularity concern, we emphasize that the ground truth underwent substantial expert verification and correction—the HMT served only as an initial automated proposal, with domain experts making the final determination for each point. Furthermore, the deep learning component introduces an independent detection stage that is not influenced by HMT; the morphological refinement is applied only within regions already identified by Faster R-CNN. To quantify the extent of independent human input into the gold standard, we compared the raw HMT output against the final expert-verified annotations across all 96 slices. Experts modified 83 points in total, i.e., 0.54% of the 15,326 ground-truth landmarks: for endpoints, 34 were added and 21 removed (the removals predominantly border truncations misclassified by the automated operator); for branch points, 28 were removed and none added. This low but non-negligible correction rate shows that the automated HMT and the expert gold standard are closely related yet not identical, so the evaluation is not a purely self-referential comparison of the HMT against itself. Nevertheless, we recognize that an ideal validation would employ independently annotated ground truth created without HMT involvement. Due to resource constraints, formal inter-rater agreement metrics (e.g., Cohen's kappa) were not computed; this represents a limitation that future work should address through more rigorous annotation protocols with multiple independent raters and quantified agreement statistics.

The complete procedure is formalized in Algorithm 1, which provides a step-by-step description of the hybrid localization pipeline.

IV. APPLICATION

This section provides a detailed explanation of each step outlined in Algorithm 1. In our study, we observed that the initial bounding box inferences generated from deep learning systems are often insufficient to capture the precise topological characteristics of complex structures such as vascular networks. Therefore, a post-processing stage is crucial to enhance the geometric precision of branching points and terminal endpoint(s).

In the first step of our refinement methodology, each

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2026.3735398

TABLE 1. Summary of Related Literature on Vascular Analysis and Landmark Detection

Reference	Year	Method	Contribution / Limitations
Zhang et al. [1]	2020	Neural network segmentation	Automated segmentation / No landmark localization
Epah et al. [2]	2018	Ultramicroscopy, micro-CT	3D quantitative analysis / Manual tuning
Goswami et al. [3]	2020	Quantitative MRA	Automated quantification / MRA only
Xu et al. [4]	2024	Traditional, DL, hybrid	Comprehensive review / No landmark detection
Villarreal et al. [5]	2024	AI + image processing	Innovative segmentation / No coordinates
Shen et al. [7]	2017	Deep learning	Comprehensive review / Not vessel-specific
Graham et al. [8]	2019	Hover-Net	Segmentation + classification / High cost
Kamnitsas et al. [9]	2017	Multi-scale 3D CNN	Multi-scale processing / Not for vessels
Zhou et al. [10]	2024	DL 3D segmentation	Bright/black blood / Box imprecision
Brugnara et al. [11]	2023	Deep learning detection	Occlusion detection / No sub-pixel
Nader et al. [12]	2023	DL classification	Bifurcation detection / Classification only
Cabrera DeBuc et al. [13]	2017	Microvascular analysis	Biomarkers / No automation
Bhuiyan et al. [18]	2012	Traditional processing	Bifurcation classification / Manual
Kaya et al. [14]	2025	Fast-rcnn, Faster R-CNN, RetinaNet, Rpn	>90% accuracy / Bounding box only
Kaya et al. [15]	2024	Faster R-CNN Hyperparameter opt.	>98% accuracy / Bounding box only
Alternative Paradigms for Landmark Localization (discussed in Section I)			
– (General)	Various	Heatmap-based keypoint regression	Probability distribution / Requires large datasets, variable morphology challenge
– (General)	Various	Dedicated keypoint heads (CenterNet, CornerNet)	End-to-end learning / Demands large annotated datasets
– (General)	Various	Skeleton-graph methods	Deterministic results / Lacks robustness to segmentation noise
Proposed	**2025**	**Faster R-CNN + HMT**	**Pixel-level accuracy (<1 px)**

Algorithm 1 Hybrid Vascular Landmark Localization

Require: Skeletonized image I, Trained model M
Ensure: Coordinates of branches B and endpoint(s) E

Stage 1: Detection
$\mathcal{D} \leftarrow M(I)$ {Detect candidates}
$\mathcal{D}_{nms} \leftarrow \text{NMS}(\mathcal{D}, 0.7)$
Stage 2: Refinement
$I_p \leftarrow \text{GlobalPad}(I, 1)$
for each $d_i \in \mathcal{D}_{nms}$ **do**
  $R_i \leftarrow \text{Extract}(I_p, d_i.bbox)$
  if $d_i.class =$ "E" **then**
    $R_i^p \leftarrow \text{Pad}(R_i, 1)$ {FG pad}
    $K \leftarrow K_E$
  else
    $R_i^p \leftarrow \text{Pad}(R_i, 0)$ {BG pad}
    $K \leftarrow K_B$
  end if
  $H_i \leftarrow \text{HMT}(R_i^p, K)$
  $(x_l, y_l) \leftarrow \text{Center}(H_i)$
  $(x_g, y_g) \leftarrow \text{ToGlobal}(x_l, y_l, d_i)$
  if $d_i.class =$ "E" **then**
    $E \leftarrow E \cup \{(x_g, y_g)\}$
  else
    $B \leftarrow B \cup \{(x_g, y_g)\}$
  end if
end for
return B, E

bounding box is designated as a Region of Interest (ROI) on the skeletonized binary image. These ROIs are small and consistently sized: across the 4,091 detected boxes on the held-out test set, the width was 5.9±0.4 px and the height 6.0±0.2 px (each side within 5–8 px), closely tracking the fixed 5×5 annotation scale. The refinement therefore operates on compact, uniform regions rather than on boxes of widely varying extent. To minimize edge artifacts arising from morphological operators, we first applied single-pixel global padding to the entire skeleton image. Then, we extracted the ROIs corresponding to the bounding boxes from the padded image. Depending on the object class (branching point 'B' or endpoint(s) 'E'), we applied a specific local padding operation to each extracted ROI. This differential padding has a strong theoretical justification. For branching points, the ROI was padded with zero-valued pixels (background). This isolates the internal pattern, ensuring that the HMT kernel, designed to detect a central pixel with multiple foreground neighbors, does not misinterpret the artificial ROI boundary as a missing connection, thus preventing false negatives at the endpoint(s). Conversely, for endpoint(s), the ROI was padded with one-valued pixels (foreground). This strategy prevents the HMT kernel, designed to detect a central pixel with a single foreground neighbor, from incorrectly identifying a true endpoint(s) at the ROI boundary as a regular point due to some of its neighbors lying outside the ROI boundary. This foreground padding effectively directs the kernel to focus on the connectivity pattern within the skeleton instead of ignoring the ROI boundary, thereby preventing false negatives. This process is illustrated in Figure 3.

After preparing the locally padded ROIs, we applied the HMT, a fundamental pattern-matching technique in mathematical morphology [19]. The HMT is particularly well-suited for this task because it operates as a precise template-matching operator that simultaneously considers both foreground and background pixel configurations. Unlike simpler morphological operations such as erosion or dilation, which

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2026.3735398

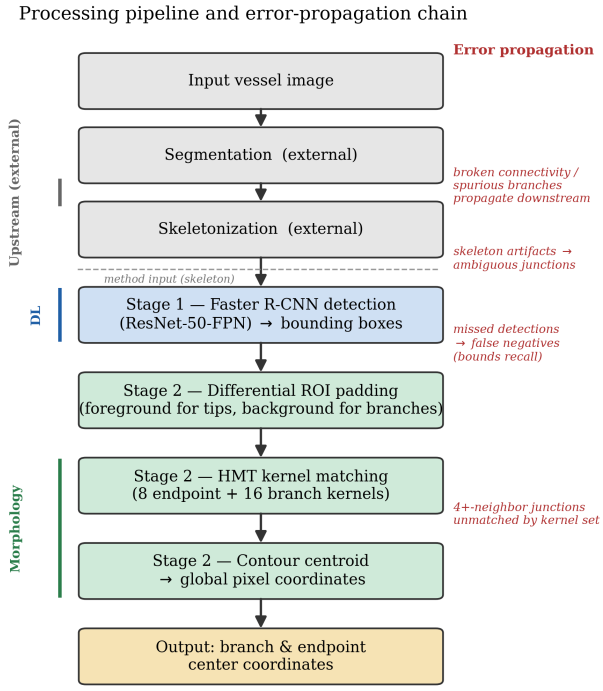


FIGURE 2. **Processing pipeline of the proposed hybrid method, made explicit as a dependency chain: upstream segmentation and skeletonization (external, quality-critical) feed the Stage 1 deep-learning detector, whose bounding boxes feed the Stage 2 morphological refinement (differential ROI padding, HMT kernel matching, contour-centroid mapping). Right-hand annotations mark where errors propagate—upstream connectivity/skeleton artifacts, detector false negatives, and junctions outside the kernel set—to the final landmark coordinates.**

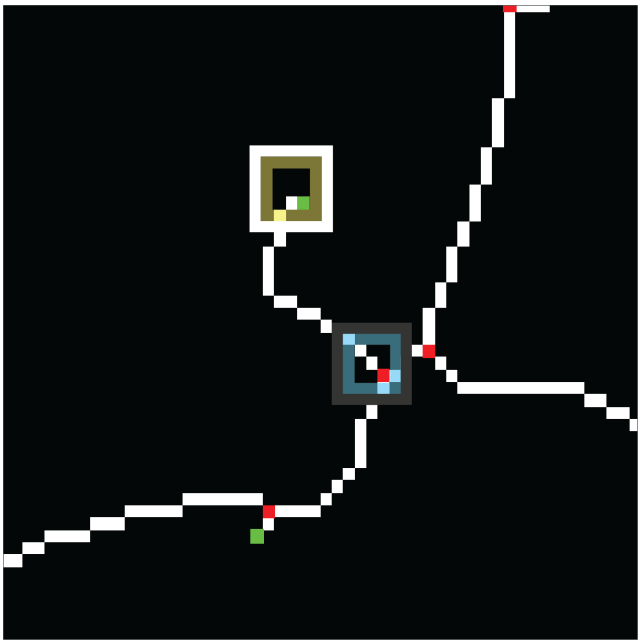


FIGURE 3. **Differential local padding of ROIs: Foreground padding (white) is applied around an endpoint(s) candidate (top-left), and background padding (black) is applied around a branching point candidate (center) to isolate the respective features.**

only examine foreground connectivity, the HMT can distinguish between topologically distinct structures by explicitly encoding the required absence of foreground pixels (background constraints) alongside presence requirements. This dual-constraint mechanism enables unambiguous discrimination between endpoints (single foreground neighbor) and branch points (three or more foreground neighbors), even when these structures appear in close spatial proximity. Furthermore, the deterministic nature of HMT ensures consistent, reproducible identification of the exact topological pixel, which is essential for achieving the sub-pixel localization accuracy that is the core objective of this study. We employed a set of morphological kernels that are designed specifically for each point type. The endpoint(s) kernels, shown in Figure 4, detect 3 x 3 patterns where a central foreground pixel has only one foreground neighbor. The branching point kernels, shown in Figure 5, detect the T- and Y-shaped patterns in which a central foreground pixel has three foreground neighbors. Active pixels in the resulting HMT response image indicate the precise locations where the morphological patterns match.

It is important to note that the HMT kernel design directly influences detection sensitivity. Our kernel sets were designed to cover all eight principal orientations (0°, 45°, 90°, etc.) of both endpoint and branching configurations on a discrete 3×3 grid. This comprehensive coverage ensures robustness to moderate structural variations in vessel morphology, such as different branching angles or vessel widths, as long as the skeleton representation maintains single-pixel-wide connectivity. The method's reliance on skeletonized input inherently normalizes vessel thickness variations, making the kernel matching insensitive to vessel diameter differences. However, severe skeletonization artifacts or highly irregular junction geometries (e.g., extremely acute-angle bifurcations) may occasionally produce ambiguous patterns that do not perfectly match any predefined kernel, representing a known limitation of template-based approaches.

We used contour detection to extract precise coordinates from the HMT response. When the HMT produces a response, we apply OpenCV's findContours function to identify connected regions of active pixels. In the majority of cases (over 98% of detections in our experiments), the HMT response within a bounding box contains exactly one connected component, making contour selection trivial. For the rare cases with multiple response regions, we select the contour whose centroid is closest to the bounding box center, as this is most likely to correspond to the detection target. We acknowledge that this heuristic, while effective in practice, is not theoretically optimal; alternative approaches such as selecting the largest response or the strongest match could be considered. The current implementation uses deterministic OpenCV contour ordering, which may vary across library

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2026.3735398

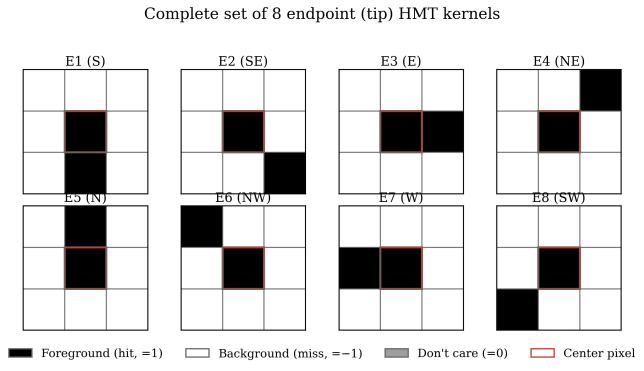


FIGURE 4. **The complete set of eight 3 × 3 morphological kernels used for endpoint(s) detection (two single-neighbor templates—one cardinal, one diagonal—each shown in its four 90° rotations). Black represents foreground, white represents background, and gray represents "do not care."**

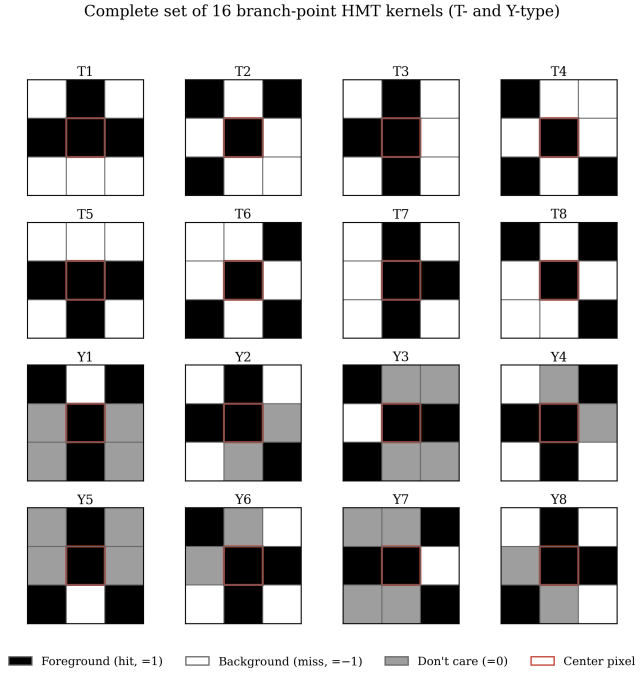


FIGURE 5. **The complete set of sixteen 3 × 3 morphological kernels used for branching-point detection: eight T-type and eight Y-type three-neighbor templates (four base templates, each shown in its four 90° rotations). Black, foreground; white, background; gray, "do not care."**

versions, though our centroid-proximity criterion provides consistent behavior regardless of ordering. The geometric center of the selected contour was considered the refined point's relative coordinates within the locally padded ROI. We then transformed these local coordinates into global image coordinates by reversing the offsets from the bounding box position and the two padding steps. This meticulous localization process eliminates geometric ambiguities from the initial deep learning detections, enabling the extraction of topological features with high accuracy.

Regarding kernel completeness: for endpoints, we employ 8 kernels covering all single-neighbor configurations at cardinal and diagonal directions. For branch points, we use 16 kernels covering the T- and Y-shaped three-neighbor junction configurations (eight of each type; see Fig. 5). While this set covers the vast majority of configurations encountered in properly skeletonized vessels, it is not mathematically exhaustive—certain rare configurations (e.g., four-or-more-neighbor junctions such as vessel crossings) may not match any kernel. In practice, such cases are extremely rare in well-preprocessed data. The complete kernel specifications are available in our code repository.

V. RESULTS

We evaluated the proposed hybrid methodology on 96 skeletonized slices from the High-Fat/High-Sucrose 13 (HFHS13) mouse brain vasculature dataset [20]. The Faster R-CNN detector was trained on 76 slices, and all quantitative metrics reported below were computed on the disjoint 20-slice held-out test set; a stricter leakage-free split is reported later in this section. Because the slices are consecutive z-sections of a single brain, their correlation structure is analyzed explicitly below. Each slice presents unique topological complexities, providing a demanding testbed for assessing the performance of our approach. A descriptive statistical analysis of the data set was conducted first to quantify key morphological characteristics, summarized in Table 2.

TABLE 2. **Descriptive Statistics of the Brain Vessel Image Dataset (N=96 slices)**

Morphological Metric	Mean	Std. Dev.	Min	Max
VAF (%)	1.1	0.37	0.32	1.83
Branching Point(s)	51.3	36.3	0	150
Endpoint(s)	108.3	55.4	3	192
Vessel Count	146.00	56.37	15	245
Total Vessel Length	3587.76	1500.00	860.82	5678.55
Avg. Vessel Length	24.70	4.89	17.43	55.95

As shown in Table 2, the dataset exhibits significant variability, with the number of branching points ranging from 0 to 150 and endpoint(s) from 3 to 192 per slice (10,400 endpoints and 4,926 branch points in total across the 96 slices). This heterogeneity confirms that the dataset provides sufficient diversity to evaluate the proposed method, while we note that all slices originate from a single brain (analyzed under leakage below).

The performance of the hybrid methodology was then quantitatively evaluated with the primary objective of assessing its accuracy in identifying branching points and endpoint(s) locations compared to the ground truth. The aggregated results are presented in Table 3. Evaluation metrics were computed based on the cumulative counts of True Positives (TP), False Positives (FP), and False Negatives (FN). A detection was considered a TP if it correctly matched a ground truth point of the same class.

Across three independent training runs, endpoint detection achieved a mean precision of 99.86%, recall of 99.15%, and

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2026.3735398

TABLE 3. Held-out Test-set Performance of the Hybrid Method (20 test slices; values are mean ± standard deviation over three independent training runs, seeds 0–2)

Metric	Endpoint(s)	Branching Point(s)
Precision (%)	99.86 ± 0.05	99.82 ± 0.06
Recall (%)	99.15 ± 0.05	94.13 ± 1.34
F1-Score (%)	99.50 ± 0.04	96.89 ± 0.73

TABLE 4. Matching-protocol sensitivity on the held-out test set (aggregate over both classes). Results are identical across thresholds 3–10 px and for greedy versus Hungarian matching.

Distance threshold	Precision (%)	Recall (%)	F1 (%)
3 px	99.8	97.1	98.5
5 px (used)	99.8	97.1	98.5
10 px	99.8	97.1	98.5

Greedy = Hungarian for every threshold tested (4, 6, 8 px omitted; identical).

F1-score of 99.50% (standard deviation ≤0.05 point), while branch-point detection achieved a mean precision of 99.82%, recall of 94.13%, and F1-score of 96.89%. The very small standard deviations—below 0.1 point for all endpoint metrics and for branch-point precision—indicate that the results are highly stable to training randomness, while the larger branch-point recall variability (±1.34) reflects the harder branch-point cases analyzed later. For a representative run (seed 0), the held-out counts were 2,129 true positives, 3 false positives, and 17 false negatives for endpoints, and 926 true positives, 2 false positives, and 74 false negatives for branch points; per-run bootstrap 95% confidence intervals (e.g., endpoint F1 [99.29–99.72], branch-point F1 [94.80–97.03]) are consistent with the across-seed variation.

We characterize variability in two complementary ways. First, the detector was retrained under three independent random seeds; the resulting held-out metrics are highly stable, and Table 3 reports the mean ± standard deviation across these runs. Second, within each run we compute bootstrap 95% confidence intervals by resampling the 20 held-out slices with replacement (2,000 iterations), which capture slice-to-slice variability. The intervals are tight for precision and endpoint recall and wider for branch-point recall, consistent with the harder branch cases (acute-angle and densely clustered junctions) discussed later. Together, the across-seed and within-run analyses characterize both the training-induced variance and the slice-to-slice sampling variability of the reported metrics.

Matching-protocol sensitivity. To confirm that the reported figures do not depend on the evaluation criterion, we swept the matching distance threshold and compared greedy against optimal (Hungarian) bipartite assignment. As summarized in Table 4, the aggregate precision, recall, and F1 are constant across thresholds from 3 to 10 pixels and identical for greedy and Hungarian matching, so the choice of matching protocol does not affect the conclusions. This invariance follows directly from the pixel-exact localization: because every true match lies essentially on its ground-truth pixel (mean error below 0.01 px), it sits far within any threshold from 3 to 10 px, and each detection has a single unambiguous nearest neighbour, so neither the distance threshold nor the assignment algorithm changes which points are matched.

Correlation structure and leakage-free evaluation. The 96 slices are consecutive z-sections of a single brain; in the original split every test slice (20 of 20) is immediately adjacent ($z \pm 1$) to a training slice. Because neighboring sections are visually similar, this adjacency could in principle inflate held-out metrics through information leakage. To measure the effect directly, we constructed a stricter contiguous block split: the highest-z block of slices (z = 81–99, 19 slices) was held out, the 76 lower-z slices (z = 3–79) were used for training, and the single boundary slice (z = 80) was dropped as a buffer so that no test slice is adjacent to any training slice (0 of 19). Retraining and re-evaluating the entire pipeline under this leakage-free protocol yielded an endpoint F1 of 99.44% and a branch-point F1 of 96.58% (Table 5, Fig. 6)—essentially identical to the original split (endpoint 99.50%, branch-point 96.89% mean over three seeds) and well within its across-seed variability. This provides direct empirical evidence that the reported accuracy is not an artifact of adjacent-slice leakage; the remaining limitation is that all data derive from a single subject.

TABLE 5. Leakage-free contiguous block split (train *z*=3–79, test *z*=81–99, buffer *z*=80 dropped; 0/19 train–test adjacency). Brackets: 95% bootstrap CIs.

Metric	Endpoint(s)	Branching Point(s)
Ground Truth Points	1446	605
True Positives (TP)	1432	565
False Positives (FP)	2	0
False Negatives (FN)	14	40
Precision (%)	99.86 [99.67–100]	100.0 [100–100]
Recall (%)	99.03 [98.57–99.49]	93.39 [91.09–95.28]
F1-Score (%)	99.44 [99.21–99.70]	96.58 [95.33–97.58]

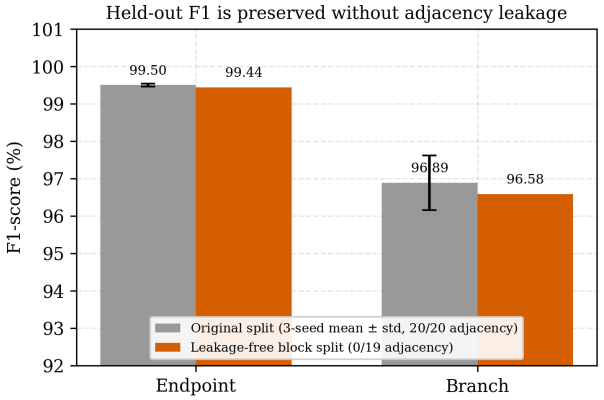


FIGURE 6. Held-out F1-score under the original random split (mean over three seeds; error bars are ± one standard deviation; 20/20 adjacent test slices) versus the leakage-free contiguous block split (single run; 0/19 adjacency). Performance is preserved when adjacency leakage is removed, indicating the reported accuracy is not driven by correlated neighboring slices.

A critical finding is the positional accuracy of **the correctly**

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2026.3735398

identified points (True Positives). For all 3,055 matched points on the held-out test set (2,129 endpoints and 926 branches, representative run), the localization error was 0 px for over 99.9% of matched endpoints and about 99.7% of matched branch points, giving a median of 0 px and a mean below 0.01 px. This near-exact agreement is an inherent consequence of the method: both the expert ground truth and the HMT output are single integer pixels lying on the one-pixel-wide skeleton, and the HMT deterministically selects the topological pixel satisfying the connectivity pattern, so a correctly matched detection typically coincides with the annotated pixel exactly. These statistics are computed only over matched true positives; localization error is undefined for missed detections (false negatives). The magnitude of the raw bounding-box-centroid error that this refinement removes—and the reason the near-zero figure above is not an independent accuracy gain—is quantified by the box-centroid baseline in the padding-ablation analysis below. This indicates that when the proposed method successfully identifies a point, it achieves pixel-exact localization (median 0 px and mean <1 px), effectively transforming imprecise bounding boxes into precise integer coordinates.

Note: The localization error is reported as <1 px because the HMT, by its morphological nature, identifies a single, centered pixel as the final point for each landmark. The term ''sub-pixel accuracy'' used elsewhere in this paper refers to mean error below one pixel; the HMT itself produces integer pixel coordinates, not fractional positions. For true positives the residual is therefore an integer number of pixels: it is zero for over 99.9% of matched detections and at most about two pixels in the rare, dense cases where the selected topological pixel differs from the annotated one.

VI. ABLATION STUDY

The central design choice in our refinement stage is the differential ROI padding applied before the HMT—foreground padding for endpoint candidates and background padding for branch candidates (Section IV). To quantify its contribution, and to complement the qualitative illustration in Fig. 3, we ablated this component on the held-out test set, comparing three settings while holding all other stages (detection, NMS, kernels, matching) fixed: (i) no ROI padding, (ii) swapped (wrong) padding, in which the endpoint and branch padding polarities are exchanged, and (iii) the proposed correct differential padding. Results are summarized in Table 6 and Fig. 7.

TABLE 6. Ablation of the differential ROI padding on the held-out test set (aggregate over both classes; mean over three seeds). "Wrong" padding swaps the endpoint/branch padding polarities.

Padding strategy	Precision (%)	Recall (%)	F1 (%)
No padding	90.42	96.81	93.50
Wrong (swapped)	92.97	68.28	78.73
Correct (proposed)	99.85	97.55	98.68

As shown in Table 6, omitting padding lowers precision to 90.4% (aggregate F1 93.5%), because the ROI boundary is misread by the HMT kernels as missing connectivity and

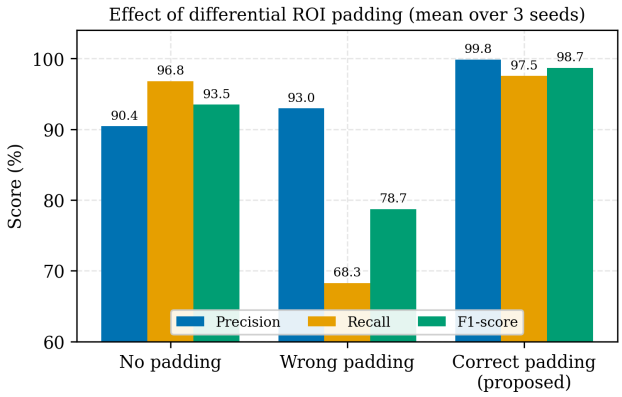


FIGURE 7. Effect of the differential ROI padding on the HMT refinement (held-out test set, aggregate over both classes). Correct class-specific padding substantially outperforms both no padding and swapped (wrong) padding, confirming that the padding polarity—not merely the presence of the HMT—drives the method's precision and F1.

produces spurious responses. Applying the wrong padding polarity is far more damaging: recall collapses to 68.3% (F1 78.7%) as many true landmarks are suppressed. The proposed differential padding recovers both, reaching 99.8% precision and 98.7% aggregate F1. This directly quantifies the effect that the differential padding illustrated in Fig. 3 has on the final result, and confirms that the class-specific padding—not merely the presence of the HMT—is essential to the method's accuracy.

This ablation isolates the padding strategy, which we identified as the most consequential design choice. To quantify the localization uncertainty that the morphological stage is designed to remove, we additionally evaluated a raw bounding-box-centroid baseline: for every matched true positive we took the Faster R-CNN box centroid directly as the landmark coordinate, without HMT refinement, and measured its Euclidean distance to the annotated point. Across the 3,138 matched detections on the held-out test set the box centroid already lies close to the annotation (mean 0.48 px, median 0.50 px, within 1 px for 96.2%), but the error is heavier-tailed for branch points (mean 0.61 px, 90.4% within 1 px) than for endpoints (mean 0.42 px, 98.9% within 1 px) and reaches up to 5 px in dense regions where neighbouring detection boxes overlap. The refinement stage removes this residual: it returns the exact one-pixel-wide skeleton pixel rather than a centroid that need not lie on the skeleton, bounds the worst-case offset, and collapses multiple overlapping detections of a single landmark to one topological point. We emphasize that the near-zero HMT localization error reported in Section V is not an independent accuracy gain over this baseline: because the reference points were themselves HMT-derived and only sparsely corrected by experts (see the Methodological Note on Ground Truth Generation), the two share an operator, so the demonstrable value of the refinement lies in guaranteeing on-skeleton integer coordinates, bounding the worst case, and de-duplicating overlapping boxes rather than in reducing the mean matched-detection error. A fully exhaustive study

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2026.3735398

would additionally vary the contour-selection rule, kernel-set size, and confidence threshold; these remain future work.

VII. DISCUSSION

Building upon the ablation in Section VI, which demonstrated that the correct differential padding raises the aggregate F1 to 98.68% (versus 93.50% with no padding and 78.73% with swapped padding), we now discuss the broader implications of these findings and analyze the error patterns in detail.

The results successfully demonstrate that our proposed hybrid approach provides a highly accurate and reliable solution for the detection and precise localization of vascular landmarks. The high held-out F1-Scores for both endpoint(s) (99.50%) and branching points (96.89%) underscore the robustness and effectiveness of the methodology.

Furthermore, the value of the deep learning stage lies in its context-aware filtering: a purely morphological approach applied to the whole skeleton flags every vessel truncation, including those at the image border, as an anatomical endpoint, producing false positives in noisy or peripheral regions, whereas the detector suppresses these artifacts, as analyzed below. This synergy between the deep learning detection stage and the morphological refinement stage—where each component compensates for the other's weaknesses—represents the core contribution of our hybrid approach.

One of the most significant findings is the notably high precision achieved for both classes (99.86% for endpoint(s) and 99.82% for branching points; see Table 3). This indicates an extremely low false positive rate, which is crucial in clinical contexts as it ensures that identified points correspond almost certainly to true anatomical features. Endpoint recall is likewise high (99.15%), whereas branching-point recall is lower (94.13%), so the residual missed detections are concentrated in the branch-point class.

A detailed error analysis reveals critical insight into this behavior. Figure 8 displays the ground truth points for a representative image generated via traditional morphological analysis. Figure 9 shows the detection results of our hybrid method on the same image.

To confirm that this qualitative behavior is representative rather than specific to a single favorable slice, Fig. 10 shows the hybrid method's detections across four slices spanning the dataset's range—from a sparse slice (z43) through a typical (z17) and a dense, cluttered slice (z64) to a border-heavy section (z57). The method continues to localize endpoints and branch points across this full range of densities, so the reported behavior is not the product of a single favorable example.

A significant portion of the false negatives are not model failures, but rather intelligent corrections to an artifact inherent in generating the ground truth. As seen in Figure 8, the traditional method incorrectly classifies any vessel segment truncated by the image border as an anatomical endpoint(s). However, our deep learning model, which uses broader contextual information from its training data, correctly identifies these truncated edges as incomplete segments and refrains

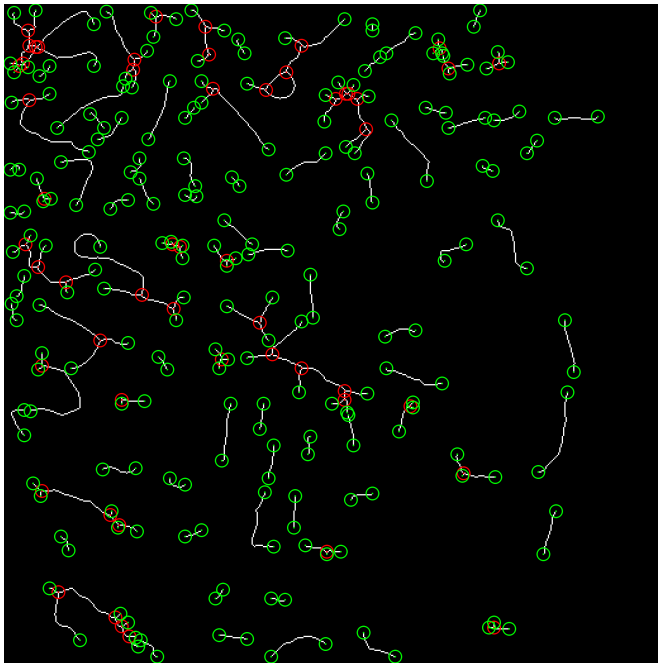


FIGURE 8. Landmarks produced by the traditional morphological (HMT) analysis on the skeletonized image HFHS13-z17. Branching points are marked in red and endpoint(s) in green, each shown as a coloured circle centred on the detected landmark point (not a bounding box). Note the endpoint(s) at the image borders where vessels are truncated.

from making false annotations (see Figure 9). This demonstrates a more context-aware understanding by the model, as it correctly avoids classifying imaging artifacts as biological features. This conservative bias improves the biological plausibility of the detections, even if it results in a lower recall score based on a flawed ground truth. Consequently, the recall values reported in Tables 3 and 6 should be interpreted as a lower bound on the model's true sensitivity to anatomically valid landmarks, because a subset of "missed" detections corresponds to artefactual labels at image borders.

Beyond the intelligent rejection of border artifacts, an analysis of the remaining false negatives revealed other contributing factors. A portion of the missed points corresponded to either extremely faint terminal vessels with low contrast or branch points where vessels merged at a very acute angle. These challenging cases represent the current limitations of the model's detection capability. Addressing these could involve targeted data augmentation with more examples of such cases or architectural enhancements, such as adding dedicated key-point detection heads.

On the held-out test set the method produced on the order of 17–19 missed endpoints and 49–74 missed branch points across the three training runs; because these totals are small, we report absolute counts rather than a sampled subset, avoiding the sampling bias inherent to reviewing only a fixed number of cases. A qualitative review of these misses indicates the recurring failure modes: for endpoints, the dominant category is vessel truncation at the image border—cases in which the

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2026.3735398

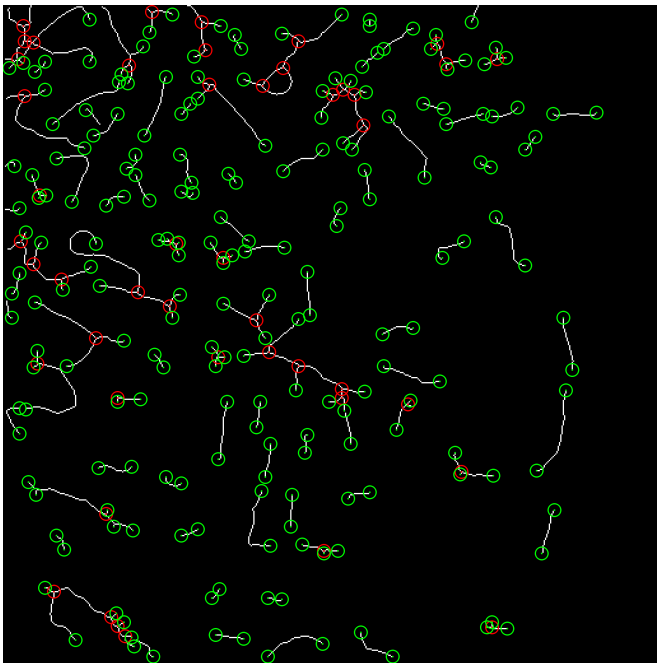


FIGURE 9. Detection results from the proposed hybrid deep learning method on image HFHS13-z17. Branching points are marked in red and endpoint(s) in green. Each detected landmark is shown as a coloured circle centred on its refined center point; these markers denote the final output coordinates, in contrast to the Stage-1 detection bounding boxes of Fig. 1. The model correctly excludes vessel terminations at the image borders but fails to detect some true internal points.

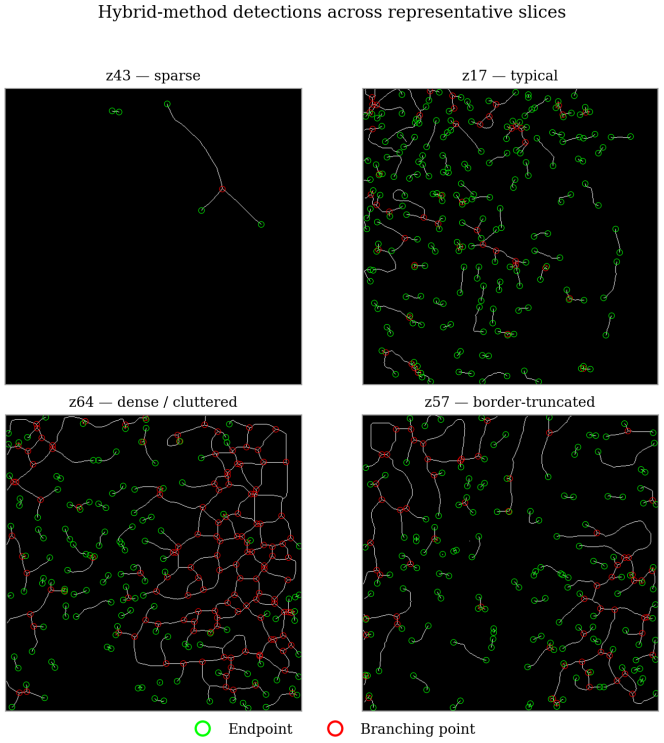


FIGURE 10. Hybrid-method detections across four slices spanning the dataset's diversity—sparse (z43), typical (z17), dense/cluttered (z64), and border-truncated (z57). Endpoint(s) in green and branching points in red, each drawn as a circle centred on the detected landmark point.

deep learning model deliberately declines to mark an artifactual termination, so that many of these ''misses'' are arguably correct rejections rather than true errors (as discussed above), with a minority corresponding to faint, low-contrast terminals in peripheral regions. For branch points, which account for the majority of the residual misses, the recurring causes are acute-angle bifurcations (vessels merging at very small angles) and dense clusters where several junctions fall within overlapping bounding boxes, together with occasional skeletonization artifacts that create ambiguous connectivity. A fully quantified per-category taxonomy with multiple independent annotators over these held-out misses would further strengthen the analysis and is left to future work.

Quantitatively, the proposed method achieves a precision above 99.7%, implying a minimal false positive rate. On the held-out test set there are only 5 false positives in the representative run (3 endpoints and 2 branch points): the endpoint false positives correspond to vessel truncations at the field-of-view boundary that were not suppressed, and the branch-point false positives arise from redundant detections at complex junctions. The very small absolute number of false positives is consistent with the near-ceiling precision reported in Table 3.

While our method focuses on precise landmark localization, recent topology-aware strategies, such as those employing centerline-based loss functions or Graph Convolutional Networks (GCNs), offer promising directions for capturing global connectivity. Additionally, recent advances in region-aware granulation-based approaches have demonstrated potential for improving spatial feature extraction in complex imaging scenarios [21], [22]. Integrating such topological constraints could further reduce false negatives in complex branching scenarios by enforcing structural continuity beyond local morphological patterns.

A key validation of our core hypothesis is the achievement of pixel-level positional accuracy for all true positive detections. The HMT-based post-processing step, by its morphological design, identifies a single, unambiguous pixel for each landmark, resulting in a localization error of less than one pixel (<1 px). This effectively resolves the inherent imprecision of bounding boxes and provides the high level of positional accuracy necessary for subsequent quantitative analyses, such as calculating segment lengths or branching angles, where precise coordinate data is an absolute prerequisite.

To make the practical significance concrete, consider two representative downstream measurements as illustrative order-of-magnitude bounds. In branching-angle estimation, the angle is computed from the directions of the daughter segments over a measurement arm of roughly 20 pixels; a coarse position estimate off by only a few pixels (on the

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2026.3735398

scale of a detection box) would induce an angular error of about $\arctan(3/20) \approx 8$–$9°$, whereas pixel-exact localization (mean <0.01 px) renders this contribution negligible. In vascular segment-length computation, a ± 3 px endpoint uncertainty on a 50-pixel segment corresponds to roughly $\pm 6\%$, whereas pixel-exact endpoints reduce this to a fraction of a percent, enabling reliable longitudinal comparisons. These estimates indicate that the pixel-level localization is not merely a technical nicety but translates into meaningfully improved reliability for the morphometric analyses that motivate the pipeline.

LIMITATIONS AND FUTURE DIRECTIONS

Despite the strong performance demonstrated, this study has several limitations that open avenues for future research.

First, and most significantly, the evaluation was conducted on 96 skeletonized slices from a single mouse brain; although the leakage-free block split (Section V) rules out adjacent-slice leakage as a confound, all slices still originate from one subject. Although the dataset exhibits significant topological heterogeneity, future work must validate the methodology on larger and more diverse datasets, including additional subjects and particularly human clinical imaging data, to ensure clinical relevance and generalizability. Specifically, the current validation is limited to a single species (mouse), a single imaging modality (light-sheet microscopy of cleared tissue), and a specific vascular bed (cerebral microvasculature). Generalization to other species (e.g., human), imaging modalities (e.g., MRA, CTA, OCT angiography), or vascular structures (e.g., retinal, coronary, or peripheral vessels) remains to be demonstrated. Differences in image resolution, contrast characteristics, noise profiles, and vessel morphology across these domains may require retraining of the deep learning component or adaptation of the HMT kernel configurations.

Second, our method assumes high-quality upstream preprocessing. The entire pipeline depends critically on accurate vessel segmentation and proper skeletonization; if these preprocessing steps introduce errors (e.g., broken connectivity, spurious branches, or incorrect topology), the subsequent landmark detection will inherit and potentially amplify these errors. We did not systematically evaluate performance degradation under varying preprocessing quality, which represents an important consideration for practical deployment.

Third, there is an inherent definition dependence in our approach. The concepts of ''endpoint'' and ''branch point'' are operationally defined through our HMT kernels based on local connectivity patterns on a discrete pixel grid. These definitions, while mathematically precise, encode a particular interpretation that may not align with all biological or clinical conventions. For instance, a structure that appears as a single bifurcation at lower resolution might resolve into multiple closely spaced junctions at higher resolution. Ground truth annotation conventions for vascular landmarks are not universally standardized, and our results should be interpreted within the context of our specific operational definitions.

Fourth, our hybrid approach combines an object detector with a separate post-processing refinement step. This study did not compare it against fully end-to-end deep learning alternatives, such as direct key-point regression models. Investigating such architectures could reveal different trade-offs in accuracy and computational efficiency and remains an important area for future comparison.

The most promising future direction involves extending this framework to 3D volumetric data from clinical imaging modalities such as MRA, CTA, and micro-CT. This extension presents both technical challenges and significant opportunities. From a deep learning perspective, the Faster R-CNN architecture would need to be replaced or adapted with 3D-capable object detectors, such as 3D Region Proposal Networks or anchor-free detectors like CenterNet3D, which can process volumetric inputs directly. The computational demands of 3D convolutions would necessitate efficient patch-based inference strategies or the use of sparse convolution techniques to maintain practical processing times.

The morphological refinement stage would similarly require adaptation. The 2D HMT kernels used in this study would be extended to 3D structuring elements capable of detecting branching configurations in volumetric space. While the number of possible 3D kernel orientations increases substantially compared to 2D, the fundamental principle of pattern matching remains applicable. The 3D extension would enable the detection of complex branching geometries, such as trifurcations and out-of-plane bifurcations, that are inherently ambiguous in 2D projections.

The clinical impact of such a 3D system would be substantial. Precise 3D coordinates of vascular landmarks could serve as direct inputs for patient-specific computational fluid dynamics (CFD) simulations, enabling the prediction of hemodynamic parameters such as wall shear stress and pressure gradients at bifurcation sites. These parameters are known to correlate with aneurysm formation and rupture risk. Furthermore, automated 3D vascular topology extraction could facilitate the development of novel imaging biomarkers for cerebrovascular diseases, including quantitative measures of collateral circulation capacity and microvascular density changes associated with small vessel disease. The integration of such a framework with existing clinical workflows could significantly enhance diagnostic capabilities and treatment planning in neurovascular medicine.

To assess generalizability, future work will include cross-dataset and cross-modality evaluations across different scanners, staining protocols, and noise/contrast regimes. We also plan a pilot validation on human MRA/CTA datasets to assess clinical transferability.

VIII. CONCLUSION

In this study, we addressed the challenge of obtaining precise coordinates for vascular landmarks from the bounding box outputs of deep learning object detectors. We proposed and validated a novel hybrid methodology that combines the contextual awareness of a deep learning model with the pixel-

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2026.3735398

level accuracy of a targeted morphological post-processing stage utilizing the HMT.

Our approach demonstrated strong performance, achieving precision rates exceeding 99.4% for both endpoint(s) and branching points. Crucially, for all correctly identified points, our post-processing refinement achieved pixel-level localization error (mean <1 px) for true positives, mitigating the positional imprecision inherent to bounding boxes. Furthermore, our analysis revealed that the deep learning model intelligently disregards imaging artifacts, such as vessel truncations at the image border, which are often misclassified by traditional methods. This demonstrates a superior, context-aware understanding that enhances the biological plausibility of the results.

We emphasize that these results were obtained on a focused evaluation of 96 skeletonized slices from a single mouse brain vasculature dataset. While encouraging, these findings represent a promising initial step rather than a definitive demonstration of general applicability. The method's performance on human clinical data, different imaging modalities, or other vascular beds remains to be established through broader validation studies. Additionally, the reported metrics depend on specific evaluation protocols (matching threshold, ground truth definition) that may influence absolute values.

In conclusion, this study presents a promising framework that demonstrates the potential to elevate deep learning object detection from a simple localization tool to a precise quantitative measurement instrument for neurovascular analysis. By successfully integrating the strengths of deep learning and traditional image processing, this work provides an initial proof-of-concept for extracting the topological features of complex vascular networks. This lays a foundation for future work toward more advanced automated analyses, including 3D network reconstruction and the development of new biomarkers for cerebrovascular diseases, pending validation on more diverse datasets and clinical populations.

To support reproducibility and facilitate further research, the source code for our implementation is publicly available at https://github.com/sametkaya/vap2d (Python 3.10, OpenCV 4.10, Detectron2 with a ResNet-50-FPN Faster R-CNN; the reported held-out metrics were produced with PyTorch 2.5.1 and CUDA 12.1, and the full dependency list with pinned versions is given in the repository `requirements.txt`). The repository includes the complete HMT kernel specifications for both endpoint and branch detection, the training/evaluation script that retrains the detector and reproduces every reported metric, confidence interval, and ablation, and the resulting model weights. Because OpenCV's contour extraction can vary across library versions and thereby affect centroid selection, the OpenCV version is pinned to ensure deterministic, reproducible behavior.

REFERENCES

[1] M. Zhang, C. Zhang, X. Wu, X. Cao, G. S. Young, H. Chen, and X. Xu, "A neural network approach to segment brain blood vessels in digital subtraction angiography," *Computer Methods and Programs in Biomedicine*, vol. 185, p. 105159, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169260719303268

[2] J. Epah, K. Pálfi, F. L. Dienst, P. F. Malacarne, R. Bremer, M. Salamon, S. Kumar, H. Jo, C. Schürmann, and R. P. Brandes, "3d imaging and quantitative analysis of vascular networks: A comparison of ultramicroscopy and micro-computed tomography," *Theranostics*, vol. 8, no. 8, pp. 2117–2133, 2018. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5928875/

[3] P. Goswami, M. K. Markey, S. J. Warach, and A. N. Dula, "Quantitative analysis of the cerebral vasculature on magnetic resonance angiography," *Scientific Reports*, vol. 10, no. 1, p. 10227, 2020. [Online]. Available: https://www.nature.com/articles/s41598-020-67225-w

[4] Y. Xu, R. Quan, W. Xu, Y. Huang, X. Chen, and F. Liu, "Advances in medical image segmentation: A comprehensive review of traditional, deep learning and hybrid approaches," *Bioengineering (Basel)*, vol. 11, no. 10, p. 1034, Oct 2024. [Online]. Available: https://www.mdpi.com/2306-5354/11/10/1034

[5] R. Villarreal, S. Chamorro-Solano, S. Cantillo, R. Pestana-Nobles, S. Arquez, Y. Vega-Sampayo, L. Pacheco-Londoño, J. Paez, N. Galan-Freyle, C. Ayala, and P. Amar, "An innovative methodology for segmenting vessel like structures using artificial intelligence and image processing," *Scientific Reports*, vol. 14, no. 1, p. 30332, 2024. [Online]. Available: https://doi.org/10.1038/s41598-024-81680-9

[6] S. T. Bhairnallykar and V. Narawade, "Exploration of image segmentation: A comprehensive review of traditional segmentation techniques," in *2023 7th International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech)*, 2023, pp. 1–6.

[7] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annual Review of Biomedical Engineering*, vol. 19, no. Volume 19, 2017, pp. 221–248, 2017. [Online]. Available: https://www.annualreviews.org/content/journals/10.1146/annurev-bioeng-071516-044442

[8] S. Graham, Q. D. Vu, S. E. A. Raza, A. Azam, Y. W. Tsang, J. T. Kwak, and N. Rajpoot, "Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images," *Medical Image Analysis*, vol. 58, p. 101563, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1361841519301045

[9] K. Kamnitsas, C. Ledig, V. F. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, and B. Glocker, "Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation," *Medical Image Analysis*, vol. 36, pp. 61–78, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1361841516301839

[10] L. Zhou, H. Wu, G. Luo, and H. Zhou, "Deep learning-based 3d cerebrovascular segmentation workflow on bright and black blood sequences magnetic resonance angiography," *Insights into Imaging*, vol. 15, no. 1, p. 81, 2024. [Online]. Available: https://doi.org/10.1186/s13244-024-01657-0

[11] G. Brugnara, M. Baumgartner, E. D. Scholze, K. Deike-Hofmann, K. Kades, J. Scherer, S. Denner, H. Meredig, A. Rastogi, M. A. Mahmutoglu, C. Ulfert, U. Neuberger, S. Schönenberger, K. Schlamp, Z. Bendella, T. Pinetz, C. Schmeel, W. Wick, P. A. Ringleb, R. Floca, M. Möhlenbruch, A. Radbruch, M. Bendszus, K. Maier-Hein, and P. Vollmuth, "Deep-learning based detection of vessel occlusions on ct-angiography in patients with suspected acute ischemic stroke," *Nature Communications*, vol. 14, no. 1, p. 4938, 2023. [Online]. Available: https://doi.org/10.1038/s41467-023-40564-8

[12] R. Nader, R. Bourcier, and F. Autrusseau, "Using deep learning for an automatic detection and classification of the vascular bifurcations along the circle of willis," *Medical Image Analysis*, vol. 89, p. 102919, Oct 2023. [Online]. Available: https://doi.org/10.1016/j.media.2023.102919

[13] D. C. DeBuc, G. M. Somfai, and A. Koller, "Retinal microvascular network alterations: potential biomarkers of cerebrovascular and neural diseases," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 312, no. 2, pp. H201–H212, Feb 2017. [Online]. Available: https://journals.physiology.org/doi/10.1152/ajpheart.00201.2016

[14] S. Kaya, B. Kiraz, and Y. Çamurcu, "Branch and end points detection in cerebral vessels images using deep learning object detection techniques," *Journal of Polytechnic*, vol. 28, no. 2, pp. 639–648, March 2025. [Online]. Available: https://doi.org/10.2339/politeknik.1492002

[15] ——, "Hyperparameter optimization in deep learning-based object detection of branching and endpoints on 2d brain vessel images," in *2024 Innovations in Intelligent Systems and Applications Conference (ASYU)*. IEEE, 2024, pp. 1–6. [Online]. Available: https://doi.org/10.1109/ASYU62119.2024.10757007

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2026.3735398

[16] J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, I. Fischer, Z. Wojna, Y. Song, S. Guadarrama, and K. Murphy, "Speed/accuracy trade-offs for modern convolutional object detectors," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 3296–3297. [Online]. Available: https://ieeexplore.ieee.org/document/8099834

[17] C. Chen, M.-Y. Liu, O. Tuzel, and J. Xiao, "R-cnn for small object detection," in *Computer Vision – ACCV 2016*, S.-H. Lai, V. Lepetit, K. Nishino, and Y. Sato, Eds. Cham: Springer International Publishing, 2017, pp. 214–230. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-54193-8_14

[18] A. Bhuiyan, B. Nath, and K. Ramamohanarao, "Detection and classification of bifurcation and branch points on retinal vascular network," in *2012 International Conference on Digital Image Computing Techniques and Applications (DICTA)*, 2012, pp. 1–8.

[19] D. Zhao and D. G. Daut, "Morphological hit-or-miss transformation for shape recognition," *Journal of Visual Communication and Image Representation*, vol. 2, no. 3, pp. 230–243, 1991. [Online]. Available: https://www.sciencedirect.com/science/article/pii/104732039190025B

[20] E. Özkan, Y. Çetin Taş, E. Şekerdağ, B. Yiğit, N. Shomalizadeh, S. Sapancı, C. Ozler, J. Kesibi, A. B. Kızılırmak, M. Pekmez, H. Yapıcı-Eser, M. Zeybel, S. Karahüseyinoğlu, and Y. Gürsoy-Özdemir, "Hyperglycemia with or without insulin resistance triggers different structural changes in brain microcirculation and perivascular matrix," *Metabolic Brain Disease*, vol. 38, no. 1, pp. 307–321, 2023. [Online]. Available: https://link.springer.com/article/10.1007/s11011-022-01100-7

[21] P. Paral, S. Ghosh, S. K. Pal, and A. Chatterjee, "Adaptive non-homogeneous granulation-aided density-based deep feature clustering for far infrared sign language images," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 9, no. 2, pp. 1269–1280, 2025.

[22] S. Ghosh, P. Paral, A. Chatterjee, and S. Munshi, "A novel 2-d robust lpp-based approach using density-based neighborhood granulation for challenging visual cue detection," *IEEE Sensors Journal*, vol. 25, no. 5, pp. 8665–8673, 2025.



SAMET KAYA received the B.S. degree in computer and control teaching from Marmara University, Istanbul, Turkey, in 2012, and a second B.S. degree in computer engineering from Sakarya University in 2015. He received the M.S. degree in computer engineering from Istanbul Aydin University in 2018, and the Ph.D. degree in computer engineering from Fatih Sultan Mehmet Vakıf University (FSMVU), Istanbul, in 2024. He is currently an Assistant Professor with the Department of Computer Engineering at Fatih Sultan Mehmet Vakıf University (FSMVU), Istanbul, Türkiye. His research interests include deep learning, object detection, natural language processing, data mining, machine learning, and image processing.



BERNA KIRAZ received the B.S. degree in Mathematics from Ege University, Izmir, Turkey, in 2004, the M.S. degree in Computer Engineering from Marmara University, Istanbul, in 2008, and the Ph.D. degree in Computer Engineering from Istanbul Technical University (ITU), Istanbul, in 2014. She was a Visiting Scholar at Michigan State University, USA, from 2014 to 2015. She served as an Assistant Professor and later as an Associate Professor in the Department of Computer Engineering and the Department of Artificial Intelligence and Data Engineering at Fatih Sultan Mehmet Vakıf University (FSMVU), Istanbul, where she also held the position of Head of the AI and Data Engineering Department. She is currently an Associate Professor in the Department of Artificial Intelligence and Data Engineering at Istanbul Technical University. Her research interests include evolutionary computation, meta-heuristics, neural architecture search, and deep learning applications in biomedical image analysis. Dr. Kiraz has published extensively in international journals such as the IEEE Journal of Biomedical and Health Informatics and IEEE Access. She is also a co-founder of Optofil, a startup developing optical imaging systems for life sciences.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/