Intracranial aneurysm detection: an object detection perspective

Youssef Assis, Liang Liao, Fabien Pierre, René Anxionnat, Erwan Kerrien

Intracranial aneurysm detection: an object detection perspective

Youssef Assis[1*], Liang Liao[1,2,3], Fabien Pierre[1], René Anxionnat[2,3], Erwan Kerrien[1]

[1] Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France.
[2] Université de Lorraine, CHRU-Nancy, Department of Diagnostic and Therapeutic Interventional Neuroradiology, F-54000 Nancy, France.
[3] Université de Lorraine, Inserm, IADI, F-54000 Nancy, France.

*Corresponding author(s). E-mail(s): youssef.assis@loria.fr;

Abstract

Note *This preprint has not undergone peer review (when applicable) or any post-submission improvements or corrections. The Version of Record of this contribution is published in " International Journal of Computer Assisted Radiology and Surgery – IJCARS 2024", and is available online at https://doi.org/10.1007/s11548-024-03132-z.*

Purpose Intracranial aneurysm detection from 3D Time-Of-Flight Magnetic Resonance Angiography images is a problem of increasing clinical importance. Recently, a streak of methods have shown promising performance by using segmentation neural networks. However, these methods may be less relevant in a clinical settings where diagnostic decisions rely on detecting objects rather than their segmentation.

Methods We introduce a 3D single-stage object detection method tailored for small object detection such as aneurysms. Our anchor-free method incorporates fast data annotation, adapted data sampling and generation to address class imbalance problem, and spherical representations for improved detection.

Results A comprehensive evaluation was conducted, comparing our method with the state-of-the-art SCPM-Net, nnDetection, and nnUNet baselines, employing two datasets comprising 402 subjects. The evaluation is based on adapted object detection metrics. Our method demonstrated comparable or even superior performance, achieving an average precision of 78.96%, sensitivity of 86.78%, and 0.53 false positives per case.

Conclusion Our method significantly reduces the detection complexity compared to existing methods, and highlights the advantages of object detection over

segmentation-based approaches for aneurysm detection. It also holds potential for application to other small object detection problems.

Keywords: Computer-aided detection, deep learning, intracranial aneurysm, magnetic resonance angiography

1 Introduction

Intracranial aneurysms are abnormal focal dilations of cerebral blood vessels. Their rupture accounts for 85% of Subarachnoid Hemorrhages (SAH), and is associated with high morbidity and mortality rates [1]. Early detection of Unruptured Intracranial Aneurysms (UIA) has become a problem of increasing clinical importance. Computed Tomography Angiography (CTA) and Magnetic Resonance Angiography (MRA) are the most commonly used imaging techniques for monitoring UIAs. Among these, 3D time-of-flight (TOF-)MRA stands out for its non-invasive nature and suitability for screening. As the number of scans performed increases each year, an automated UIAs detection method would be valuable to assist radiologists in their clinical routine.

1.1 Related works

Aneurysm detection, or more broadly object detection is a specific image analysis task that differs from image classification and segmentation. While image classification categorizes the entire image based on aneurysm presence, image segmentation assigns labels to voxels for precise aneurysms delineation. In contrast, aneurysm detection identifies individual aneurysms location and assigns confidence scores to each.

Early Computer-Aided Detection (CAD) systems for aneurysm detection relied on predefined imaging features, and their real-world performance have not been fully established [2, 3]. With the advent of deep learning, Convolutional Neural Networks (CNNs) became the standard for various visual tasks. However, UIAs detection remains challenging due to multiple factors such as limited labeled datasets, scarce aneurysm signals in MRA data, absence of prior knowledge regarding aneurysm number and locations, and significant computational demands for 3D image processing. Several 2D deep learning methods have been proposed for aneurysm detection [4–6]. However, recent approaches have shifted toward fully 3D approach, utilizing contextual information from the three anatomical planes via patch-based techniques. Although the dual-path multi-scale DeepMedic model was deemed promising [1], the high false positives remains a challenge [7]. To address this, a two-stage approach was introduced in [8], combining Regions-of-Interest (RoIs) identification with precise aneurysm segmentation. Nevertheless, most existing works have been evaluated on small, private datasets, limiting their reproducibility and multi-site analyses.

In 2020, the Aneurysm Detection And SegMentation (ADAM) challenge [9] was conducted to objectively compare various aneurysm detection methods. Most methods relied on the UNet segmentation architecture [10]. To address class imbalance, loss functions and data augmentation were employed. Various data sampling approaches

were observed, with methods based on large contextual patches ranking highest on the leaderboard. However, these patches require significant computational power for training and extensive datasets for reliable sample modeling. The choice of UNet was primarily due to the complexity of object detection approaches, which involve multiple loss functions, advanced sampling strategies, and anchor definitions [11]. Yet, segmentation-based methods may not be the most efficient for object detection tasks. In such methods, like nnUNet [12], post-processing steps (e.g., thresholding, connected component extraction and analysis) are required to locate objects, and voxel-based probabilities do not integrate into confidence scores at object level. These post-processing steps can negatively impact performance compared to object detection approaches that incorporate these steps during the learning process [13]. Additionally, the discrepancy between ground truth object description and predicted voxel maps leads to less appropriate object-level evaluation metrics [11]. In contrast, traditional object detection approaches are assessed using the Average Precision (AP) metric, considering overlap with ground truth, at different confidence thresholds [11, 14, 15].

The leading method in the ADAM challenge 2020, nnDetection [11], stands out as the only method following an object detection approach. It relies on the anchor-based RetinaUNet [13] for detecting aneurysms using 3D bounding boxes. Anchor-based methods, such as RetinaNet [16] and YOLO [15], have become standard in natural image object detection, with RetinaNet offering higher accuracy but slower speed compared to YOLO, which offers a better balance between speed and accuracy. However, these anchor-based methods introduce model complexity and flexibility limitations, requiring customized anchor hyper-parameters for each dataset and creating class imbalance between positive and negative anchors [17]. This issue has led to the emergence of anchor-free detectors, characterized by simpler designs, fewer parameters, and potential performance improvements [14]. Additionally, bounding box representation may not be ideal for spherical objects such as aneurysms [14]. Existing works have investigated circle [14] and ellipse [18] representations for 2D images. In the realm of 3D biomedical objects, such alternative representations remain limited. The SCPM-Net model [19], for instance, uses a spherical representation to detect lung nodules in CT scans and utilizes an anchor-free method with a loss function that optimizes the Intersection-over-Union (IoU) between spheres. However, it incorporates multiple hyperparameters that need to be adapted to each problem. Another challenge in current aneurysm detection works is the limited datasets, mainly due to labor-intensive voxel-wise annotation. In 2022, a public aneurysm TOF-MRA dataset was released, combining voxel and weak labels [20].

1.2 Contributions

In this work, we introduce an anchor-free single-stage method for UIAs detection. Our approach is built upon a spherical representation of aneurysms, which simplifies and accelerates the annotation process. To mitigate limited training data, we adopt a strategy involving small patches, along with data sampling and synthesis techniques, effectively addressing class imbalance and aneurysm scarcity in images. We validate the efficacy of our approach through an ablation study on a private dataset. Cross-validation experiments were performed using a public and large dataset, which was

meticulously curated by two experts to ensure reliable ground truth. Our method demonstrates state-of-the-art performance compared to the baselines, highlighting the advantages of the spherical representation over standard bounding boxes for UIAs.

2 Materials and methods

2.1 Datasets and annotation

Two TOF-MRA aneurysm datasets were used. The first dataset consists of 132 TOF-MRA exams (75 female, 57 male) collected at our medical institution between 2015 and 2021 according to the following inclusion criteria: diagnosed unruptured saccular aneurysms smaller than 20 mm, no pre-treated aneurysm. A single exam was included per patient. All images were acquired using a 3T scanner. Acquisition parameters included TR = 28 ms, TE = 3.4 ms, slice thickness = 0.8 mm, FOV = 24, flip angle = 17°, 4 slabs (54 slices/slab), and acquisition time = 6min 28s, resulting in $512 \times 512 \times 254$ volumes with a $0.47 \times 0.47 \times 0.4$ mm^3 voxel size. The dataset contained a total of 206 aneurysms (mean size: 3.97 ± 2.32 mm), ranging from 1 (84/132) to 5 aneurysms (4 subjects) per image. Of these, 81 aneurysms were smaller than 3 mm, and 77 measured between 3 and 5 mm.

The second dataset is the public dataset [20], which consists of 296 images. After applying the same inclusion criteria as the in-house dataset, a total of 270 images were kept for analysis. Previous works on aneurysm detection relied on voxel-wise labeling, which is time-consuming and susceptible to intra- and inter-rater variability. To address these issues and make annotating large datasets much easier, weak labels approximating aneurysms with spheres have recently been investigated, achieving comparable performance to voxel-wise labeling [20, 21]. Our method [21], however, provides greater annotation precision by defining spheres using two points: the center of the neck and the dome, instead of the oversized spheres used in [20], which may include healthy blood vessels (see Fig. 1). An expert interventional neuroradiologist with 14 years of experience (LL) used our method to annotate the in-house dataset, and review and correct the public dataset [20]. During this correction step, 7 new aneurysms were added, and 5 aneurysms were removed from the provided annotations as they were considered simple dysplasia or irregularities on the vessel surface. The images of dataset [20] contained a total of 164 aneurysms (mean size: 3.74 ± 2.17 mm), ranging from 0 (130 healthy subjects) to 5 aneurysms (4 subjects) per image.

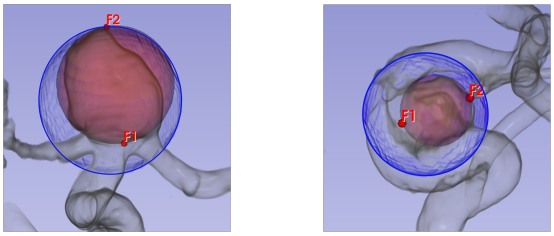


Fig. 1 Annotation comparison in 3D volume rendering: oversized annotation (blue), including surrounding healthy vessels [20], versus our proposed annotation (red) to better approximate aneurysms.

Of these, 66 aneurysms were smaller than 3 mm, and 72 between 3 and 5 mm. These annotations are referred to as *initial annotations* in the paper.

2.2 Data selection and generation

The discriminative power of an object detector is a measure of its capacity to accurately model both objects and the background. In large patch approaches, an aneurysm is present in the majority of patches, which requires the dataset to contain healthy subjects for reliable background modeling. Our method employs small patches, allowing for the extraction of multiple non-overlapping negative, aneurysm-free patches. These patches are $96 \times 96 \times 96$ voxels in size, with an isotropic voxel size of 0.4 mm. However, this approach introduces class imbalance, with only one positive patch for each aneurysm. To tackle this, we used adapted data sampling strategies. The first strategy involved duplicating positive patches 50 times and applying random distortions to diversify aneurysm shapes [21]. Another class imbalance arises from vascular information, comprising about 3 to 5% of the background signal. To distinguish healthy blood vessels from aneurysms, our second strategy involved pre-selecting 40 non-overlapping negative patches for each patient data. Out of these, 30 were centered on blood vessels, identified using the brightest voxels as patch centers. Each training epoch included all available positive patches, supplemented with randomly selected negative patches, constituting 15% of the training set. We also applied data augmentation techniques, such as random rotations (0 to 180°), translations (0 to 10 mm), and horizontal flips.

2.3 Network architecture

We propose a 3D anchor-free approach adapted for small object detection using a spherical representation. Following a grid-based approach, our architecture divides the input patch of $96 \times 96 \times 96$ voxels into $12 \times 12 \times 12 = 1728$ cells (see Fig. 2). Each cell is responsible for detecting a single object, based on the presence or absence of

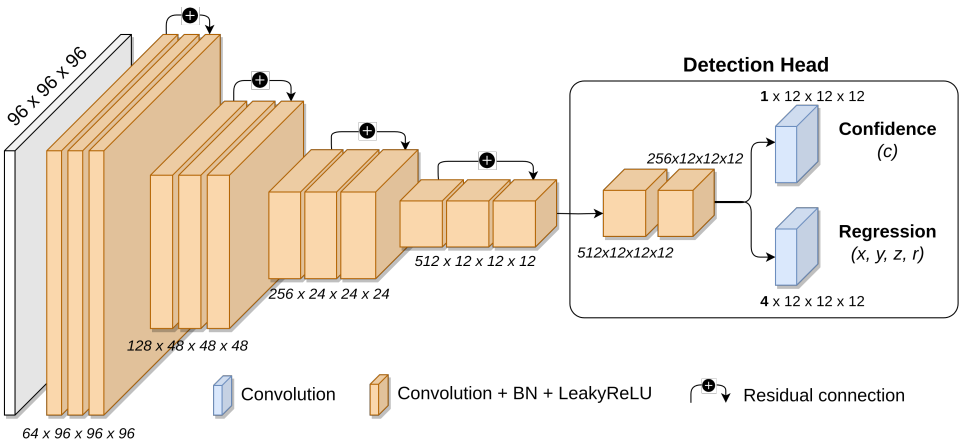


Fig. 2 Our proposed CNN : object detection using spheres (x, y, z, r) and confidence scores (c). 2D feature maps are used only for visualization purposes; their actual sizes in 3D are displayed.

its center within the cell. Each detection corresponds to a sphere centered on the aneurysm center (x, y, z) with a radius (r) equal to half the aneurysm size.

The input patch is encoded into feature maps through a series of residual convolutional blocks. The *Detection Head* further processes these feature maps by dividing them into a grid of 1728 cells using two parallel convolutions. The first convolution yields a confidence probability score (c) indicating the presence of an aneurysm center within each cell. For positive cells (i.e., containing an aneurysm center), the second convolution estimates the aneurysm center $(x,\ y,\ z)$ and radius (r). However, this approach introduces a high imbalance between negative (N) and positive (P) cells. To address this, and inspired by 2D YOLO [15], we adopted a weighted loss function that combines terms for detection confidence, localization, and size estimation. For confidence scores, we computed binary cross-entropy (BCE) using class probabilities for both positive (c_P) and negative (c_N) cells. We then weighted the negative term by half the number of positive cells $(\#P)$ in the training batch (see Eq. 1). For more precise optimization of aneurysm localization $(x,\ y,\ z)$ and size (r) on such small objects, we applied the mean squared error (MSE) for positive cells, with a weight factor of 5 to account for their scarcity (see Eq. 2).

$$\text{Classification } (P, N) = \text{BCE } (c_P) + 0.5 \times \#P \times \text{BCE } (c_N) \tag{1}$$

$$\text{Regression } (P) = 5 \times \sum_{i=1}^{\#P} \text{MSE } (x_i, y_i, z_i, r_i) \tag{2}$$

2.4 Detection and performance evaluation

Our method was implemented using PyTorch with stochastic gradient descent (SGD) and the following hyperparameters: 200 epochs, balanced batch sampling, a batch size of 32, and an initial learning rate of 10^{-2}. Hyperparameters were chosen after preliminary experiments on a subset of the in-house dataset. During inference, we used a patch reconstruction approach. The original volume is divided into $96 \times 96 \times 96$ voxel patches with a voxel resolution of 0.4 mm, allowing for a 16 voxel overlap to address border effects from convolutions. Detections were generated from each patch and retained if their center $(x,\ y,\ z)$ fell within the central $64 \times 64 \times 64$ region of the patch. To eliminate overlapping detections, Non-Maximum Suppression (NMS) was applied with an IoU threshold of 10%. Both training and inference were performed using a NVIDIA RTX A6000 GPU.

The evaluation of object detection methods requires metrics that considers both localization and confidence scores. Existing aneurysm detection works often use metrics less adapted for object detection tasks. For instance, UNet-based approaches define objects as connected components (CC) of thresholded output probability maps. A paradoxical consequence is that a higher confidence threshold can increase the number of CC and hence detections. Similarly, average output probabilities favor smaller objects around the probability peak. In our work, we evaluate the methods with metrics based on the IoU between the predicted and ground truth (GT) spheres at an IoU threshold of 10% [11, 22]. A GT sphere with an IoU score above 10% was tagged as a true positive (TP), else it was a false negative (FN). A false positive (FP) was counted

for each predicted sphere with no IoU score above 10%. This leads to the competition of the average precision ($AP_{0.1}$) score. We also report sensitivity ($Sensitivity_{0.5}$) and false positives per case ($FPs/case_{0.5}$) at a default 50% confidence threshold.

We compared our method against SCPM-Net [19], nDetection [11], and nnUNet [12], with the latter two ranking first and third in the ADAM Challenge 2020 [9]. SCPM-Net also employs $96 \times 96 \times 96$ voxels patches, but its loss function and random training patch selection impeded model convergence, leading us to adopt our own data strategy without optimizing IoU between spheres. On the other hand, nnDetection and nnUNet are self-configuring methods, eliminating manual setup. nnDetection employs anchor-based object detection with 3D bounding boxes, and 27 predefined anchors. nnUNet specializes in medical image segmentation, utilizing a modified UNet architecture [10] and deep supervision training. Both baselines employ large patches with diverse data augmentation techniques. For a fair comparison, we converted the outputs of nnDetection and nnUNet to spherical representations. For nnDetection, we transformed predicted bounding boxes into spheres using the largest extent of the box as the sphere diameter. For nnUNet, we fitted one sphere per connected component in the segmented voxel image. The sphere diameter was determined as the maximum distance between two voxel locations, and the confidence score was set as the maximum predicted voxel value.

3 Experiments and Results

3.1 Ablation study

To evaluate the contribution of each component of our method, we conducted an ablation study using the in-house dataset. The dataset was split into two subsets: a training set with 92 subjects (138 aneurysms, average size of 4.11 ± 2.48 mm) and a test set with 40 subjects (68 aneurysms, average size of 3.67 ± 1.93 mm). We compared *Model0* (our method) with 6 model variants. *Model0* was trained with 8096 patches, including 1196 negatives. The 6 model variants were:

- ***Model1***: Trained without applying random distortion to positive patches.
- ***Model2***: Trained exclusively with positive patches, omitting negative ones.
- ***Model3***: Trained using all available candidate negative patches, making up 34% of the total training patches (as compared to the 15 % used in *Model0*).
- ***Model4***: Trained using oversized ground truth spheres (as in [20], 33% increase in diameter compared to *Model0*).
- ***Model5***: Trained using bounding boxes (x, y, z, width, height, and depth) instead of spherical ground truths.
- ***Model6***: Follows an anchor-based approach with 6 anchor spheres ranging in diameter from 2.31 mm to 19.63 mm, determined by k-means on the training set.

As shown in Table 1 (top), our experiments demonstrate the superior performance of our proposed method. *Model1* showed that removing random distortion from positive patches led to reduced detection confidence scores, indicating the importance of diverse aneurysm shapes for accurate identification. Unlike traditional 2D object detectors that primarily rely on positive images, *Model2* highlights the importance

Models	$AP_{0.1}$ (%)	$Sensitivity_{0.5}$ (%)	$FPs/case_{0.5}$
Model0	80.24	77.94	0.40
Model1	74.69	72.06	0.35
Model2	71.72	70.59	0.57
Model3	77.66	77.94	0.60
Model4	69.59	70.59	0.37
Model5	72.47	75.00	0.72
Model6	81.64	70.59	0.37
Our method	81.22 ± 5.34	81.55 ± 5.40	0.57 ± 0.06
SCPM-Net [19]	80.14 ± 6.49	82.03 ± 7.50	0.71 ± 0.14
nnDetection [11]	80.88 ± 6.33	77.15 ± 8.84	0.68 ± 0.13
nnUNet [12]	79.45 ± 4.70	76.04 ± 4.19	0.27 ± 0.10

Table 1 Models comparison on the in-house dataset: our proposed method is *Model0*. (Top) Ablation study models; (Bottom) Baseline comparisons, presented as mean ± standard deviation scores across 5 cross-validation test subsets.

of negative patches in our 3D aneurysm detection approach, particularly in improving background modeling. *Model3* shows that this background modeling is related to an increased sensitivity. However, too many negative samples adversely affects foreground accuracy, leading to more FPs. *Model4* found that using oversized ground truth spheres detrimentally affected detection performance. Employing tighter-fitting spheres improves the discrimination between healthy vessels and aneurysms. *Model5* validated the superiority of spherical representation over classical bounding boxes for detecting aneurysms, which are ball-shaped objects. In particular, bounding boxes tend to generate more false positives. Finally, *Model6* demonstrated that adopting an anchor-based approach is unnecessary for UIAs detection using spheres, as the sphere radius can be efficiently optimized during training. Pre-defined anchors exacerbate the imbalance between positive and negative cells, leading to lower confidence scores.

To further assess the performance of our method and measure variability in different data instances, we conducted a 5-fold cross-validation on the in-house dataset (132 subjects). The dataset was divided into 5 subsets, with 4 (105 or 106 subjects) for training and 1 subset (27 or 26 subjects) for testing. The performance results of our method and baselines are presented in Table 1 (bottom). Our method achieved an average precision of 81.22%, sensitivity of 81.55%, and 0.57 FPs/case. Different training and inference times were measured. Each model of our proposed method required 24 hours for the training phase and 30 minutes for inference, against, respectively, 35 hours and 1 hour 24 minutes for nnDetection, 28 hours and 1 hour for nnUNet, and 32 hours and 30 minutes for SCPM-Net.

3.2 5-Fold cross-validation

To evaluate the robustness of our approach against state-of-the-art baselines, we conducted a 5-fold cross-validation study on the public dataset [20]. This dataset comprises 270 subjects, including 130 healthy subjects, from various medical institutions, reflecting real-world scenarios. It was divided into five subsets of 54 subjects each, maintaining a balanced mix of health and pathology subjects. In each fold, our method used an average of 7875 training patches, including 1075 negative ones.

Evaluation	Methods	$AP_{0.1}$ (%)	$Sensitivity_{0.5}$ (%)	$FPs/case_{0.5}$
Initial annotations	Our method	75.80 ± 7.11	84.76 ± 5.00	0.57 ± 0.17
	SCPM-Net	72.41 ± 7.89	85.36 ± 6.18	0.75 ± 0.19
	nnDetection	73.68 ± 6.38	84.76 ± 4.72	0.67 ± 0.12
	nnUNet	72.46 ± 4.74	71.95 ± 9.11	0.13 ± 0.06
Revised annotations	Our method	78.96 ± 5.24	86.78 ± 3.98	0.53 ± 0.14
	SCPM-Net	76.35 ± 5.91	86.20 ± 5.22	0.71 ± 0.15
	nnDetection	78.01 ± 6.07	86.78 ± 6.16	0.63 ± 0.10
	nnUNet	73.47 ± 4.17	71.26 ± 8.09	0.11 ± 0.04

Table 2 5-Fold cross-validation on the public dataset [20]. All methods were trained using the *initial annotations*. Evaluation scores are reported as mean ± standard deviation over 5 test subsets, with *initial* (top) and revised (bottom) annotations.

Table 2 (top) compares the performance of our method with the baselines. Our method achieved the highest average precision score of 75.80%, with a sensitivity score of 84.76% and 0.57 FPs/case. Following a visual review of the detections, we found that many unlabeled aneurysms were detected, raising concerns about the reference annotation quality. Therefore, a new, more robust ground truth was established. Two interventional neuroradiology experts (LL and RA) with 14 and 32 years of experience separately conducted a blinded review of detections for each method, including FN annotations, resulting in three annotations per expert (one for each method). Decisions for each detection were made by majority vote, with consensus used to resolve ties (7.6% of detections). The review revealed that 23 FPs were actually aneurysms initially missed during the initial annotation (see Fig. 5a), while 13 aneurysm annotations were, in fact, irregularities on the vessel surface (see Fig. 5b). Therefore, the total number of aneurysms increased from 164 to 174, with a mean size of 3.65 ± 2.15 mm. We then re-evaluated the performance of the methods based on the revised annotations.

As shown in Table 2 (bottom), all methods demonstrated improved performance. Our method achieved an average precision of 78.96%, sensitivity of 86.78%, and 0.53 FPs/case, exhibiting comparable or even superior performance to baseline methods. In Figure 3 (left and middle), the Precision-Recall and FROC curves are presented. Figure 3 (right) demonstrates performance analysis of our method based on aneurysm size. For small aneurysms (< 3 mm), our method achieved 77.22% sensitivity, while SCPM-Net, nnDetection, and nnUNet reached 75.94%, 73.42%, and 45.57%, respectively. For medium-sized aneurysms ($3 - 5$ mm), our method exhibited a sensitivity of 98.59%, compared to 97.18%, 98.59%, and 92.96% for SCPM-Net, nnDetection,

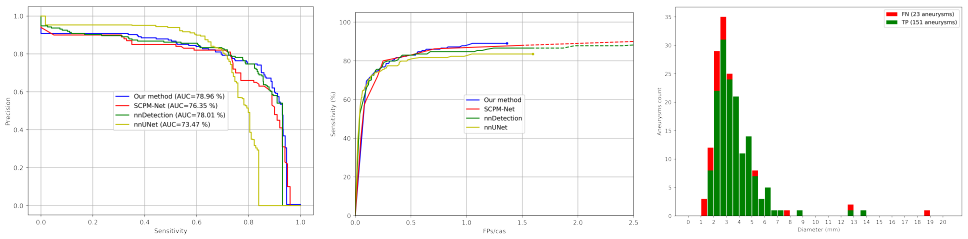


Fig. 3 Precision-Recall (left) and FROC (middle) curves for the methods, with SCPM-Net and nnDetection showed higher false positives. (Right) Detection rate per aneurysm size for our method.

and nnUNet, respectively. For larger aneurysms (> 5 mm), our method demonstrated a sensitivity of 83.33%, whereas SCPM-Net, nnDetection, and nnUNet achieved sensitivities of 87.5%, 95.83%, and 91.67%, respectively.

4 Discussion

Despite recent promising results, aneurysm detection remains a significant challenge. Previous works have proved difficult to evaluate and compare due to the lack of common datasets and appropriate metrics. In response to this, we have introduced a 3D anchor-free object detector adapted for small objects like aneurysms. Our method considers spherical representations, enabling a fast and straightforward annotation. It relies on small, non-intersecting negative patches, and data synthesis for positive patches, expanding the number of available training samples for a better leverage of small datasets. The evaluation was performed on large and public dataset using IoU-based metrics, involving the assessment of each detection by two experts for robust ground truth. Our method performed better than nnUNet and similar, even slightly better than nnDetection, achieving an average precision of 78.96%, sensitivity of 86.78%, and 0.53 FPs/case.

Our model architecture, inspired by YOLO, performs competitively compared to larger architectures like RetinaUNet [13]. It involves fewer parameters and hyperparameters, avoiding computationally intensive large patches. Our experiments favor tight-fitting spheres for aneurysm annotation over oversized spheres, as those adopted in dataset [20] (see Fig. 1). Spherical representation outperforms standard bounding boxes, as it requires only 4 parameters (x, y, z, radius) compared to 6 (x, y, z, width, height, depth) for boxes, simplifying training. Spheres provide rotation invariance, aligning with rotation-based data augmentation (see Fig. 4).

To address the high class imbalance, we used data synthesis for positive patches and sampling strategies for negative patches, improving discrimination between aneurysms and healthy vessels and cerebral tissue. We acknowledge that, as a downside of the small patch approach, large aneurysms (> 10 mm) are more challenging to detect for our method. It should also be noted that a very small sample of only 4 aneurysms was available for that category. Moreover, radiologists tend to excel in detecting larger aneurysms, making automated assistance more crucial for smaller ones. Challenging false positives often occur at small arteries branching. Due to resolution limitations and slower flow, the MRI signal rapidly gets closer to the noise floor, which let our

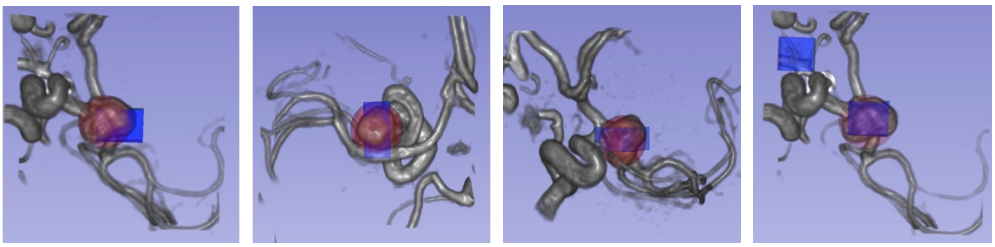
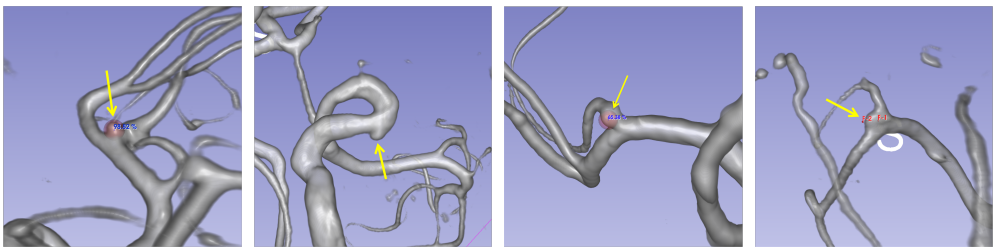


Fig. 4 Rotation invariance: Box (blue) and spherical (red) representations. Instability of box size relative to aneurysm volume. A FP in box representation appears (right).



(a) (b) (c) (d)

Fig. 5 Detection spheres (red) and their confidence scores (blue). (a) Unlabelled aneurysm of 2.4 mm detected by our method. (b) False annotation at the branching of a small artery. (c) FP detection (irregularity at the middle cerebral artery bifurcation). (d) FN aneurysm of 1.91 mm.

method mistake such cases for small aneurysms (see Fig. 5c). Out of the 174 aneurysms in the dataset [20], 14 of the 38 FN aneurysms had sizes below 2.5 mm (see Fig. 5d). Part of these misses are related to the annotation uncertainty on such aneurysms, which are difficult to diagnose in TOF-MRA. It is important to note that the models were trained using initial (noisy) annotations and evaluated by two clinical experts.

The new ground truth based on a thorough assessment by the experts allows for an evaluation of the initial annotations of dataset [20]. It shows 82.7% sensitivity and 0.07 FPs/case. This is on par with previously reported findings that clinical experts have a good sensitivity and an excellent specificity. If current deep learning methods are able to exceed the clinicians' sensitivity, reaching their specificity requires further effort. Detection performance is particularly poor for small aneurysms, which should therefore be a priority in future works.

5 Conclusion

In this paper, we presented a novel object detection approach for aneurysm detection, using a simple-to-do spherical representation and efficient small patch data strategy. Our extensive evaluation, conducted on a large public dataset by two experts, highlights annotation uncertainty, a challenge we aim to tackle in future research. Despite this, our method was compared against three state-of-the-art approaches: SCPM-Net and nnDetection, following an object detection paradigm, and nnUNet, based on segmentation. Our first result is a further hint that object detection approaches are better suited than segmentation-based methods to detect aneurysms. Second, our method performs similarly or better than SCPM-Net and nnDetection with lower computational cost, making it promising for other small object detection problems.

Statements and Declarations

Acknowledgments. This work was co-funded by the Grand-Est Region and the University Hospital of Nancy, France.

Code availability. Our code and both initial and revised annotations used for dataset [20] are available at : https://gitlab.inria.fr/yassis/DeepAneDet.

Conflict of interest. The authors declare that they have no conflict of interest.

Ethical approval. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed consent. Informed consent was obtained from all individual participants included in the study.

References

[1] Sichtermann, T., Faron, A., Sijben, R., Teichert, N., Freiherr, J., Wiesmann, M.: Deep learning–based detection of intracranial aneurysms in 3D TOF-MRA. American Journal of Neuroradiology **40**(1), 25–32 (2019)

[2] Arimura, H., Li, Q., Korogi, Y., Hirai, T., Abe, H., Yamashita, Y., Katsuragawa, S., Ikeda, R., Doi, K.: Automated computerized scheme for detection of unruptured intracranial aneurysms in three-dimensional magnetic resonance angiography. Academic radiology **11**(10), 1093–1104 (2004)

[3] Yang, X., Blezek, D.J., Cheng, L.T., Ryan, W.J., Kallmes, D.F., Erickson, B.J.: Computer-aided detection of intracranial aneurysms in MR angiography. Journal of digital imaging **24**(1), 86–95 (2011)

[4] 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. Journal of Magnetic Resonance Imaging **47**(4), 948–953 (2018)

[5] Ueda, D., Yamamoto, A., Nishimori, M., Shimono, T., Doishita, S., Shimazaki, A., Katayama, Y., Fukumoto, S., Choppin, A., Shimahara, Y.: Deep learning for MR angiography: automated detection of cerebral aneurysms. Radiology **290**(1), 187–194 (2019)

[6] Stember, J.N., Chang, P., Stember, D.M., Liu, M., Grinband, J., Filippi, C.G., Meyers, P., Jambawalikar, S.: Convolutional neural networks for the detection and measurement of cerebral aneurysms on magnetic resonance angiography. Journal of digital imaging **32**(5), 808–815 (2019)

[7] Faron, A., Sichtermann, T., Teichert, N., Luetkens, J.A., Keulers, A., Nikoubashman, O., Freiherr, J., Mpotsaris, A., Wiesmann, M.: Performance of a deep-learning neural network to detect intracranial aneurysms from 3d tof-mra compared to human readers. Clinical neuroradiology **30**, 591–598 (2020)

[8] Joo, B., Ahn, S.S., Yoon, P.H., Bae, S., Sohn, B., Lee, Y.E., Bae, J.H., Park, M.S., Choi, H.S., Lee, S.-K.: A deep learning algorithm may automate intracranial aneurysm detection on mr angiography with high diagnostic performance.

European Radiology **30**(11), 5785–5793 (2020)

[9] Timmins, K.M., Schaaf, I.C., Bennink, E., Ruigrok, Y.M., An, X., Baumgartner, M., Bourdon, P., De Feo, R., Di Noto, T., Dubost, F.: Comparing methods of detecting and segmenting unruptured intracranial aneurysms on tof-mras: The adam challenge. Neuroimage **238**, 118216 (2021)

[10] Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3d U-Net: learning dense volumetric segmentation from sparse annotation. In: International Conference on Medical Image Computing and Computer-assisted Intervention, pp. 424–432 (2016). Springer

[11] Baumgartner, M., Jäger, P.F., Isensee, F., Maier-Hein, K.H.: nndetection: A self-configuring method for medical object detection. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 530–539 (2021). Springer

[12] Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods **18**(2), 203–211 (2021)

[13] Jaeger, P.F., Kohl, S.A., Bickelhaupt, S., Isensee, F., Kuder, T.A., Schlemmer, H.-P., Maier-Hein, K.H.: Retina U-Net: Embarrassingly simple exploitation of segmentation supervision for medical object detection. In: Machine Learning for Health Workshop, pp. 171–183 (2020). PMLR

[14] Nguyen, E.H., Yang, H., Deng, R., Lu, Y., Zhu, Z., Roland, J.T., Lu, L., Landman, B.A., Fogo, A.B., Huo, Y.: Circle Representation for Medical Object Detection. IEEE transactions on medical imaging **41**(3), 746–754 (2021)

[15] Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: Unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788 (2016)

[16] Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2980–2988 (2017)

[17] Yi, J., Wu, P., Liu, B., Huang, Q., Qu, H., Metaxas, D.: Oriented object detection in aerial images with box boundary-aware vectors. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 2150–2159 (2021)

[18] Chen, J., Zhang, Y., Wang, J., Zhou, X., He, Y., Zhang, T.: Ellipsenet: Anchor-free ellipse detection for automatic cardiac biometrics in fetal echocardiography. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 218–227 (2021). Springer

[19] Luo, X., Song, T., Wang, G., Chen, J., Chen, Y., Li, K., Metaxas, D.N., Zhang, S.: Scpm-net: An anchor-free 3d lung nodule detection network using sphere representation and center points matching. Medical image analysis **75**, 102287 (2022)

[20] Di Noto, T., Marie, G., Tourbier, S., Alemán-Gómez, Y., Esteban, O., Saliou, G., Cuadra, M.B., Hagmann, P., Richiardi, J.: Towards automated brain aneurysm detection in TOF-MRA: open data, weak labels, and anatomical knowledge. Neuroinformatics, 1–14 (2022)

[21] Assis, Y., Liao, L., Pierre, F., Anxionnat, R., Kerrien, E.: An efficient data strategy for the detection of brain aneurysms from MRA with deep learning. In: Deep Generative Models, and Data Augmentation, Labelling, and Imperfections, pp. 226–234 (2021). Springer

[22] Assis, Y., Liao, L., Pierre, F., Anxionnat, R., Kerrien, E.: Aneurysm pose estimation with deep learning. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2023, pp. 543–553. Springer, Cham (2023)