Vessel-aware aneurysm detection using multi-scale deformable 3D attention

Alberto M. Ceballos-Arroyo[1], Hieu T. Nguyen[1], Fangrui Zhu[1], Shrikanth M. Yadav[2], Jisoo Kim[2], Lei Qin[2], Geoffrey Young[2], and Huaizu Jiang[1]

[1] Northeastern University
ceballosarroyo.a,nguyen.trungh,zhu.fang,h.jiang@{northeastern.edu}
[2] Brigham and Women's Hospital
smyadav@mgh,jkim@bwh,lei_qin@dfci,gsyoung@bwh{.harvard.edu}

Abstract. Manual detection of intracranial aneurysms (IAs) in computed tomography (CT) scans is a complex, time-consuming task even for expert clinicians, and automating the process is no less challenging. Critical difficulties associated with detecting aneurysms include their small (yet varied) size compared to scans and a high potential for false positive (FP) predictions. To address these issues, we propose a 3D, multi-scale neural architecture that detects aneurysms via a deformable attention mechanism that operates on vessel distance maps derived from vessel segmentations and 3D features extracted from the layers of a convolutional network. Likewise, we reformulate aneurysm segmentation as bounding cuboid prediction using binary cross entropy and three localization losses (location, size, IoU). Given three validation sets comprised of 152/138/38 CT scans and containing 126/101/58 aneurysms, we achieved a Sensitivity of 91.3%/97.0%/74.1% @ FP rates 0.53/0.56/0.87, with Sensitivity around 80% on small aneurysms. Manual inspection of outputs by experts showed our model only tends to miss aneurysms located in unusual locations. Code and model weights are available online.

Keywords: Aneurysm detection · CT angiography · Deep learning

1 Introduction

Intracranial aneurysms (IAs) are instances of bulging cerebral blood vessels, with ruptures accounting for around 85% of subarachnoid hemorrhages. They are related to high mortality rates (23-50%) [7] and a considerable (10-20%) risk of permanent disability. To find and diagnose IAs, clinicians review 3D CTA imaging datasets comprised of hundreds of slices per patient. This process is time-consuming and exhausting, increasing the risk of missing potentially deadly IAs. Thus, the design of automatic solutions for this task has gained increased clinical interest. Most such solutions are segmentation-[4,25,30,31] or detection-based [1,29]; nevertheless, both usually suffer from critical difficulties such as: 1) **a high number of false positives (FPs) per scan**, averaging around 3 FPs per scan for models with Sensitivity $\geq$ 80% (Fig. 1); and 2) **much lower sensitivity on smaller IAs**, ranging from 30% to 60% [6,26].

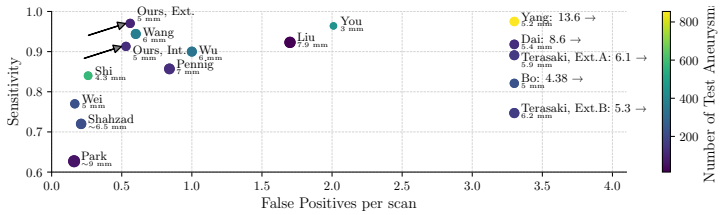

Fig. 1: Lesion-level Sensitivity vs FP rate, we compare our results on the test partitions of a public dataset [4] with recent IA detection work [5,12,17,19,20,23,25,26,27,29,30]. We display test set size (encoded by color) and mean aneurysm diameter. Note: all studies (except Bo) depicted use private datasets, precluding a direct comparison. *Adapted from Bizjak and Špiclin* [3].

High FP rates can be linked to the need to run models on large 3D scans, leaving more room for errors. As aneurysms are located in blood vessels, researchers have tried to reduce FPs by enhancing balanced sampling [6,31] with **vessel information**. Specifically, they have used vessel segmentations as a **hard constraint** on sampling when training and choosing the areas to run inference on [8]. With accurate vessel segmentations, training/inference can be done on vessels and nearby areas, reducing potential FPs; however, this can require costly imaging techniques like dynamic CTA. The alternative is to use neural networks or morphological methods, but due to the challenging nature of the problem [10] these are less likely to segment all vessels accurately, causing models to fail to detect aneurysms in non-segmented areas when a hard constraint is applied.

The comparatively lower detection performance on smaller aneurysms in recent work [6,23,25,29,30] could be associated with the low resolution of deeper layers in 3D neural networks. In the broader computer vision literature, small object detection is addressed via aggregating **multi-scale** features extracted from different layers of models like CNNs [11]. However, 3D data is denser, leading to larger computational costs that can complicate the use of multi-scale features, especially for models with compute-intensive attention mechanisms [24].

Another key problem in automatic aneurysm detection research is evaluation [3]. Namely, there is **limited consistency** in case-control design and metrics choice; in some studies, no evaluation is done on external datasets [15] nor across aneurysm sizes [17]; in others, the evaluation metrics are not comprehensive, complicating assessments of the trade-off between Sensitivity and FP rate [3]; and often neither code nor data are made available, rendering most proposals irreproducible and difficult to compare with, save a few exceptions [4,6,29]. There is thus a pressing need to use comprehensive *and* reproducible training *and* evaluation pipelines that allow to better assess the impact of ongoing research.

In this work, we tackle aneurysm identification with a model comprised of a convolutional 3D encoder and a lightweight detection Transformer [24]. We first feed 3D patches into the encoder to produce multi-scale features which are enhanced with a 3D positional encoding combining spatial and scale informa-

tion with a **distance map** describing how far each voxel is from the closest vessel. The Transformer detects aneurysms via a **multi-scale deformable 3D attention** mechanism [32] that allows it to integrate information across scales before making predictions. Thanks to these design choices, our model can exploit multi-scale details to achieve high Sensitivity at a low FP rate (see Fig. 1); and, unlike the approaches that use vessel information as a hard constraint, ours integrates vessel proximity in a **soft** manner by making it part of the positional encoding, meaning the model can detect aneurysms even where the vessel information is not accurate. Moreover, we compute metrics patient- and lesion-wise across aneurysm sizes. Since our results are mostly based on a public dataset comprising data from various hospitals, they are likely to be more **reproducible** and generalizable. We will also release our code and model weights to the entire research community. Our contributions can be summarized as follows:

- We adapt multi-scale deformable attention to 3D aneurysm detection, with a state-of-the-art (SOTA) sensitivity of +90% overall and around 80% for small aneurysms, with few FPs (under 0.6 per scan in average)
- We leverage vessel distance maps to inform the model about the location of segmented vessels in a soft way without outright discarding patches that do not intersect with them, reducing the potential for missed aneurysms
- We implement a comprehensive, reproducible evaluation pipeline and report metrics on several cross-institutional datasets and aneurysm sizes while accounting for healthy and ill patients

2 Materials and methods

2.1 Data

We use the public intra-cranial aneurysm segmentation dataset published by Bo *et al.* [4]. This dataset consists of an internal training partition ($1,186$ cases/$1,363$ aneurysms), one internal validation set (152 cases/126 aneurysms), and two external validation sets (138 cases/101 aneurysms in total). Additionally, we collected our own private test partition of 38 scans/58 aneurysms annotated as bounding volumes by a radiologist using an internal annotation tool. We resample scans and segmentation masks to 0.4 mm spacing with **SimpleITK** [14]. Since our work centers on intra-cranial aneurysms, we crop scans to exclude most of the torso and neck while making sure to keep all aneurysms from the validation/testing sets. Finally, we transform each aneurysm mask in the public data by computing its minimum bounding cuboid's *height*, *width*, *depth*, and center coordinates $c = (x, y, z)$. For each scan, we segment the vessels with a 3D U-Net, as part of a concurrent study, trained on dynamic-CTA-based [18] ground truth segmentations using **nnU-Net** [9]. Upon training the segmentation model, we validated it on 11 CT scans annotated semi-automatically by a radiologist, achieving a 0.929 modified Dice Coefficient ([Pred $\cap$ GT]/GT). Based on the segmentation results, we used the signed distance transform [21] to compute the

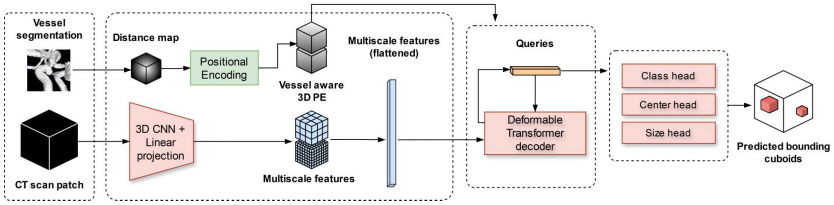

Fig. 2: The proposed architecture.

distance (in voxels) from each non-vessel voxel to the nearest vessel voxel and vice-versa, thus producing a *vessel distance map*.

For *sampling*, we use patch size 64^3 with spacing 0.4 mm as specified above, resulting in patches with a side length of 25.6 mm. To account for the imbalance of patches with and without aneurysms, we sample on average 50% of patches containing aneurysms and the rest from areas without them. During training, we augment each patch randomly with flipping (on every axis), translation (0-20 voxels on every axis), rotation (0-20 degrees in the z axis), and spacing transforms (90%-120% of the base spacing). Intensity values in Hounsfield units are clipped to the range $[0, 800]$ and then normalized to be in $[-1, 1]$.

2.2 Model architecture

Our detection architecture (Fig. 2) is comprised of two modules: the first is a 3D CNN from which we extract feature maps at multiple resolutions, and the second is a decoder-only Transformer based on the one-stage Deformable DETR [32].

For the *convolutional encoder*, we employ a 3D CNN architecture with squeeze-and-excite-based channel attention. Specifically, we use the layers of CPM-Net [22] up to the bottleneck block to extract up to four levels of 3D feature maps at 1/2, 1/4, 1/8, and 1/16 of the spatial resolution of the input volume x. Given these multi-scale input features, which we deem $x_{\triangle}$, we use a trainable linear layer $\triangle$ to map them all to the same feature space d_m. Next, we flatten the features into tokens $x_{\underline{\triangle}}$ and feed them into the decoder's *multi-scale deformable attention module* together with a small set of n_q trainable reference points (queries) Q of size d_m. Deformable attention allows each query q to dynamically attend to n_k points (or keys) around them across scales [32].

The coordinates of the n_k keys are determined with another trainable linear layer O with weights W_o and biases b_o, which we use to compute offsets relative to the location of each query. After computing the locations each query will attend to, we use tri-linear interpolation to gather key features K from the input features $x_{\underline{\triangle}}$. Given Q and K, we compute self-attention among the queries followed by cross-attention from the queries to the keys to obtain new representations for each query. This process can be repeated for an arbitrary number of layers.

To encode *positional and vessel information*, we use an absolute sinusoidal positional encoding (PE) [16]. We compute PE_l for each level of the multi-scale features; to keep distance representations consistent across levels coordinates

(x, y, z) are scaled to $[0, 1]$. Vessel awareness is achieved by using vessel distance as a 4th spatial dimension in the positional encoding. As the distance maps contain un-normalized values, we divide them by the patch size. Scale information for each level is encoded in a learnable embedding p_l that is added to PE_l. In line with DETR literature, we use the PE in the attention mechanism by adding it to the query/key features before computing the attention scores; thus, positional *and* vessel information gets injected into the model once per layer. On top of making the positional encoding three-dimensional, adapting the attention mechanism to 3D required us to initialize offset biases b_o accordingly by sampling uniformly from the unit sphere (as opposed to the unit circle in the 2D case). Zhu *et al.* [32] also processes the multi-scale features with a Transformer encoder, treating every token as a query. However, we found that doing so made convergence slower on top of the extra computational load. Other elements, such as skip connections and layer normalization modules, are used as described in [32].

After the Transformer module, we feed the queries into 3 multilayer perceptrons for classification, center localization, and cuboid size prediction. If there are aneurysms in the input patch, Hungarian matching is used to pair each detection with either an aneurysm or the background. To facilitate convergence, given a GT aneurysm, we always match it to the closest detection (irrespective of distance) to compute the loss. Other detections are only matched to a GT if the L2 distance between their centers is under 0.5 voxels (around 0.2 mm). Our *loss function* $\mathcal{L}$ is a weighted sum of 4 losses: binary cross entropy (BCE), L2 center loss, L2 size loss, and bounding cuboid IoU loss. The BCE loss is always computed, while the rest are only computed for detections matched to an aneurysm. *Inference* is done by running the model on overlapping patches from the entire scan using a sliding window with size 64^3, stride 32. We remove redundant predictions using non-maximum suppression with a small threshold of $t_{nms} = 0.05$, as we run inference on overlapping 3D crops and would like to avoid having multiple detections for the same aneurysm.

2.3 Evaluation metrics

We define a **positive detection** based on confidence score threshold t_+ chosen by considering the performance on the *training* partition. As for the definition of a **successful detection**, prediction $\hat{a}$ is a true positive (TP) of ground truth (GT) aneurysm a if their intersection over union (IoU) surpasses threshold $t_{IoU} = 0.3$, selected by showing radiologists 3D visualizations of aneurysm detections at various thresholds; we depict how the chosen t_{IoU} looks in Fig. S1. For the private data, we use intersection over the minimum (IoM) to account for the less tight annotations compared to bounding cuboids derived from the public data's segmentation masks. For **segmentation model evaluation**, we use the metric defined in [4]; given the predicted segmentation for a scan, we extract all connected components as detections $\hat{a}$. For each pair of aneurysm a, and possible detections $\hat{a}$, we compute centers c, $\hat{c}$ and radii r, $\hat{r}$, and $\hat{a}$ is deemed a TP detection of a if $dist(c, \hat{c}) \leq (r + \hat{r})$. We make this metric fairer for the baseline segmentation model by defining the radius $\hat{r}$ of a segmented aneurysm as 1/2 of

Table 1: Detection performance @ $t_{IoU} = 0.3$ (t_{IoM} for the private data), with t_+ for each model chosen based on Se@FPr=1 using the train data. We report Se@FPr curves to illustrate the threshold-agnostic performance of each model in Fig. S2, as well as results @ $t_{IoU} = 0.5$ in Tab. T1.

			Se ↑ (%)						
Dataset	**Model**	t_+	**All**	**Small**	**Med.**	**Large**	**P-Se** ↑	**FPr** ↓	**P-Sp** ↑
Pub. In.	PARQ	.9	85.7	65.5	92.2	85.7	84/102	0.47	**38/50**
	nnDet.	.6	81.7	75.8	83.3	85.7	86/102	1.20	11/50
	CPM-Net	.8	88.8	72.4	**95.5**	71.4	88/102	**0.42**	32/50
	Ours	.95	**91.3**	**82.8**	93.3	**100**	**91/102**	0.53	36/50
Pub. Ex.	PARQ	.9	83.2	64.3	84.9	92.9	75/92	0.59	31/46
	nnDet.	.6	90.1	64.3	93.1	**100**	85/92	1.37	10/46
	CPM-Net	.8	87.1	64.3	97.3	57.1%	79/92	**0.44**	**35/46**
	Ours	.95	**97.0**	**78.6**	**100**	**100**	**89/92**	0.56	**35/46**
Private	nnDet.	.6	62.1	N/A	N/A	N/A	22/38	0.73	N/A
	CPM-Net	.8	65.5	N/A	N/A	N/A	22/38	**0.63**	N/A
	Ours	.95	**74.1**	N/A	N/A	N/A	**26/38**	0.87	N/A

the longest side of its minimum bounding cuboid. We do this since the baseline outputs detailed segmentation masks while our model outputs bounding cubes that overestimate aneurysm sizes, which would make for an unfair comparison.

Lesion-level Sensitivity (Se) is defined as the fraction of aneurysms that are detected by the model. The **False Positive rate (FPr)** is defined as the mean number of FP predictions per scan. Lesion-level **Se@FPr** measures the Sensitivity of a model at a given FPr, also computed across scans. **Patient-level Sensitivity (P-Se)** evaluates how often the model detects *all* aneurysms in a scan containing them, while **Patient-level Specificity (P-Sp)** evaluates how often the model makes *zero* FP predictions for a healthy patient [3]. We do **Size-based evaluation** by computing **Se** & **FPr** across aneurysm sizes, based on diameter: *small* (0-3 mm), *medium* (3-7 mm), *large* (over 7 mm). Bo *et al.* [4] do not provide aneurysm sizes, so we approximate them to ellipsoids to derive diameter d from the annotation's volume as in [3,19], thus: $d = \rho \cdot \sqrt[3]{\frac{3V}{4\pi}}$, with $\rho = 1.45$. Using this criterion, we group aneurysms in the public dataset as follows: 29 small/90 med./7 large (internal); 14 small/73 med./14 large (external). Due to the nature of our private dataset's annotations, it was not possible for us to compute size-based metrics on said partition.

3 Experiments and results

We train all detection models on the internal training split of the public data and we use the testing splits for evaluation. Models were trained using **PyTorch 2.0.1** [16] on a node with 4 NVIDIA A5000 GPUs (24GB VRAM), 48 CPUs, and 256 GB of RAM. We optimized the model for 68,000 steps with AdamW [13], using a cosine learning rate scheduler going from $lr = 10 \times e^{-4}$ to $5 \times e^{-6}$.

Table 2: Comparison with *GLIA-Net* in the segmentation to detection setting.

Dataset	**Model**	t_+	**Se** ↑	**P-Se** ↑	**FPr** ↓	**P-Sp** ↑
Pub. In.	GLIA-Net	–	83.3%	82/102	4.27	1/50
	Ours	.95	**90.5%**	**90/102**	**0.375**	**38/50**
Pub. Ex.	GLIA-Net	–	72.3%	68/92	5.22	1/46
	Ours	.95	**96.0%**	**88/92**	**0.56**	**35/46**

For sampling/augmentation, we followed Sec. 2.1 with batch size 64, sampling 8 patches from 8 different scans per batch. We used input size 64^3, with multiscale feature sizes $(32^3, 16^3, 8^3, 4^3)$. For the decoder, we used 2 layers with $d_m = 384$ and $n_q = 8$ queries with $n_h = 16$ attention heads, each allowed to attend to $n_k = 32$ offset points for each of the $n_l = 4$ feature levels. We weigh the losses as follows: $w_{\mathcal{L}_{BCE}} = 4$, $w_{\mathcal{L}_{IoU}} = 1$, $w_{\mathcal{L}_{center}} = 0.1$, and $w_{\mathcal{L}_{size}} = 0.1$.

Quantitative results: in Table 1, we compare with 3 detection baselines: A) *PARQ*, a 3D recurrent detection Transformer [28]; B) *CPM-Net*, a one-stage, 3D detection CNN model first proposed for lung nodule detection [22]; and C) *nnDetection*, a SOTA self-configuring pipeline for medical image detection based on the Retina-U-Net architecture [2] which obtained the first place in the 2020 ADAM challenge for IA detection. We trained A) and B) using patch size 64^3 and our sampling/augmentation pipeline; for *nnDetection* we ran the authors' automated pipeline, which produced a 3-model ensemble. In Table 2, we use the approach described in Section 2.3 to compare with *GLIA-Net* [4], a coarse-to-fine U-Net-based segmentation model. Since *GLIA-Net* was trained on the same training dataset as ours, we use the model checkpoints released by the authors.

Compared with PARQ, nnDetection, and CPM-Net, our approach achieves higher Sensitivity by a margin of 5-15% across all aneurysm sizes, with a major improvement for small IAs in particular. Moreover, it retains a FPr that is competitive (within a margin of 0.25 FPs per image) with CPM-Net and PARQ while resulting in around half the FPs as the SOTA nnDetection approach. Our segmentation, EDT, and detection pipeline takes 7 min. in a workstation with a RTX 3090 GPU, 32 CPU cores, and SSD storage. Since the target scan processing time before human review is 20 min., this time is considered acceptable.

Ablation study: In Table 3, we report the impact of various choices, including how vessel information is exploited (either softly or as a hard post-processing threshold), the type of attention employed, and the use of data augmentation. We find that using vessel information in a soft rather than hard way helps the model better generalize to the external data, with a minor Sensitivitiy-FPr trade-off reflected in the choice of using it as part of the *PE* or as an additional input channel. Overall, there is a significant reduction in the FPr from using vessel information at multiple scales. Likewise, we note an statistically significant 10-20% performance boost with multi-scale deformable attention vs. single-scale dense attention, highlighting its usefulness in detecting varying-size aneurysms.

Table 3: Ablation study. $t_+ = 0.95$, $t_{IoU} = 0.3$ for all models. Changes with respect to model ① in *italics*. We measure significance w.r.t. ① via permutation testing of the case-wise Sensitivity and number of FPs, between parentheses (* indicates statistically significant result with $p < 0.05$).

		Attn. mech.			Pub. In. [4]		Pub. Ex. [4]	
V.	**Vess. aw.**	**Scale**	**Type**	**Aug**	**Se %** ↑ (p)	**FPr** (p) ↓	**Se %** ↑ (p)	**FPr** ↓ (p)
①	Pos. enc.	Mult	Deform	✓	**91.3**	0.53	**97.0**	0.56
②	Pos. enc.	Mult	Deform	✗	6.35 (*)	1.91 (*)	3.96 (*)	2.47 (*)
③	Pos. enc.	*Mult* $(8^3, 4^3)$	Deform	✓	**91.3** (1.00)	0.51 (0.89)	94.1 (0.66)	0.72 (0.17)
④	Pos. enc.	*Single* (4^3)	Deform	✓	86.5 (0.69)	0.71 (*)	92.1 (0.67)	0.87 (*)
⑤	*Input ch.*	Mult	Deform	✓	90.5 (1.00)	**0.32** (*)	93.1 (0.75)	**0.51** (0.66)
⑥	*No*	Mult	Deform	✓	**91.3** (1.00)	0.76 (*)	96.0 (1.00)	1.05 (*)
⑦	*Post-proc.*	Mult	Deform	✓	86.5 (0.69)	0.75 (*)	92.1 (0.67)	1.04 (*)
⑧	*No*	*Single* (16^3)	*Dense*	✓	87.3 (0.75)	0.84 (*)	77.2 (*)	1.71 (*)

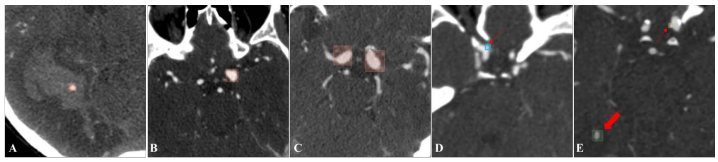

Fig. 3: TPs (A-C), FN (D), and FPs (E). **A.** Right distal PCA aneurysm. **B.** Broad-based left supra-clinoid ICA aneurysm. **C.** Bilateral fusiform aneurysm of the supra-clinoid ICAs. **D.** Tiny right para-clinoid aneurysm. **E.** Atherosclerotic calcification of left ICA (thin arrow). Choroid plexus in the right lateral ventricle (thick arrow). *(PCA: posterior cerebral artery; ICA: internal carotid artery).*

We must note that multi-scale dense attention caused OOM errors due to the input size (37k tokens), so we only tested dense attention on a single 16^3-sized feature map (4k tokens). The worst results correspond to using no augmentation; this is expected as the training pipeline is run only on 1,363 unique aneurysms, so the model overfits and is unable to generalize neither to the test data.

4 Discussion and Conclusion

We proposed a novel approach for intracranial aneurysm detection in CTA scans using multi-scale deformable 3D attention and auxiliary vessel segmentations. To the best of our knowledge, our model achieves state-of-the-art results on the largest public intracranial aneurysm dataset [4] and beats 3 strong detection baselines encompassing CNN- and Transformer-based detection models. Compared with other methods in the literature (see Fig. 1 and Table 2), ours has consistently higher Sensitivity while incurring an average of 0.55 FPs per scan. Despite a drop in performance on small aneurysms, our model surpasses all baselines by a margin of 10-15% Sensitivity for that size class. Extensive experiments (Table 3) suggest that injecting vessel information several times as part of the

positional encoding may work better than using it as a hard constraint [8] or as an extra input modality [27]; and that using multi-scale approaches is key to achieve more consistent performance across aneurysm scales. Although poor segmentations could harm detection performance, our results on several datasets indicate vessel awareness improves/conserves detection Sensitivity.

As shown in Fig. 3 our model successfully detects and generates tight bounding cuboids for IAs of various types, *e.g.* saccular (incl. broad-based) and fusiform (Fig. 3. A, B, C). Some undetected aneurysms include a tiny aneurysm at a turning point of the para-clinoid ICA (Fig. 3. D), and those in unusual locations such as the cortical branch of the middle cerebral artery or callosomarginal branch of the anterior cerebral artery. Importantly, these IAs would also be difficult for human experts to detect. Interesting FPs include atherosclerotic calcifications, which can mimic IAs even clinically due to the resultant vessel wall irregularity (Fig 3. E). One of the model's 14 FNs (see Fig. S3) in the public test data resulted from detecting 2 IAs in the same bounding cuboid; clinicians reviewing the outputs deemed this case not likely to be an usability problems. Out of the 13 other FNs, 8 would have been detected using less strict values of $t_{IoU} = 0.1$ and $t_{+} = 0.8$, while remaining under 1.5 FPs per scan. Our evaluation pipeline proved helpful for clinical collaborators to understand the pros and cons of our work, and the derived insights motivate us to explore its clinical applicability.

Acknowledgments. The authors acknowledge the financial support provided by NIH grant 1R01LM013891-01A1. Ceballos-Arroyo, A. is grateful for the funding provided by Colombia's Minciencias and Fulbright under the Fulbright Minciencias 2021 program.

Disclosure of Interests. The authors declare no competing interests.

References

1. Assis, Y., Liao, L., Pierre, F., et al.: Aneurysm Pose Estimation with Deep Learning. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2023, vol. 14221, pp. 543–553. Springer Nature Switzerland, Cham (2023)
2. Baumgartner, M., Jäger, P.F., Isensee, F., Maier-Hein, K.H.: nnDetection: A Self-configuring Method for Medical Object Detection. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2021. pp. 530–539. Lecture Notes in Computer Science, Springer International Publishing, Cham (2021)
3. Bizjak, Ž., Špiclin, Ž.: A Systematic Review of Deep-Learning Methods for Intracranial Aneurysm Detection in CT Angiography. Biomedicines **11**(11), 2921 (Nov 2023). https://doi.org/10.3390/biomedicines11112921
4. Bo, Z.H., Qiao, H., Tian, C., et al.: Toward human intervention-free clinical diagnosis of intracranial aneurysm via deep neural network. Patterns **2**(2), 100197 (Feb 2021). https://doi.org/10.1016/j.patter.2020.100197
5. Dai, X., Huang, L., Qian, Y., et al.: Deep learning for automated cerebral aneurysm detection on computed tomography images. International Journal of Computer Assisted Radiology and Surgery **15**(4), 715–723 (Apr 2020)
6. Di Noto, T., Marie, G., Tourbier, S., et al.: Towards Automated Brain Aneurysm Detection in TOF-MRA: Open Data, Weak Labels, and Anatomical Knowledge. Neuroinformatics **21**(1), 21–34 (Jan 2023)

7. Etminan, N., Chang, H.S., Hackenberg, K., et al.: Worldwide Incidence of Aneurysmal Subarachnoid Hemorrhage According to Region, Time Period, Blood Pressure, and Smoking Prevalence in the Population: A Systematic Review and Meta-analysis. JAMA Neurology **76**(5), 588–597 (05 2019)
8. Ham, S., Seo, J., Yun, J., et al.: Automated detection of intracranial aneurysms using skeleton-based 3D patches, semantic segmentation, and auxiliary classification for overcoming data imbalance in brain TOF-MRA. Scientific Reports **13**(1), 12018 (Jul 2023). https://doi.org/10.1038/s41598-023-38586-9
9. Isensee, F., Jaeger, P.F., Kohl, S.A.A., et al.: nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. Nature Methods **18**(2), 203–211 (Feb 2021). https://doi.org/10.1038/s41592-020-01008-z
10. Kim, B., Oh, Y., Wood, B.J., et al.: C-DARL: Contrastive diffusion adversarial representation learning for label-free blood vessel segmentation. Medical Image Analysis **91**, 103022 (Jan 2024). https://doi.org/10.1016/j.media.2023.103022
11. Lin, T.Y., Dollar, P., Girshick, R., et al.: Feature Pyramid Networks for Object Detection. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 936–944. IEEE, Honolulu, HI (Jul 2017)
12. Liu, X., Mao, J., Sun, N., et al.: Deep Learning for Detection of Intracranial Aneurysms from Computed Tomography Angiography Images. Journal of Digital Imaging **36**(1), 114–123 (Feb 2023)
13. Loshchilov, I., Hutter, F.: Decoupled Weight Decay Regularization (Jan 2019). https://doi.org/10.48550/arXiv.1711.05101
14. Lowekamp, B., Chen, D., Ibanez, L., Blezek, D.: The Design of SimpleITK. Frontiers in Neuroinformatics **7** (2013). https://doi.org/10.3389/fninf.2013.00045
15. Park, A., Chute, C., Rajpurkar, P.e.a.: Deep Learning–Assisted Diagnosis of Cerebral Aneurysms Using the HeadXNet Model. JAMA Network Open **2**(6), e195600 (Jun 2019). https://doi.org/10.1001/jamanetworkopen.2019.5600
16. Paszke, A., Gross, S., Massa, F., et al.: Pytorch: an imperative style, high-performance deep learning library. In: Proceedings of the 33rd International Conference on Neural Information Processing Systems (2019)
17. Pennig, L., Hoyer, U.C.I., Krauskopf, A., et al.: Deep learning assistance increases the detection sensitivity of radiologists for secondary intracranial aneurysms in subarachnoid hemorrhage. Neuroradiology **63**(12), 1985–1994 (Dec 2021)
18. Rajiah, P.S., Weber, N., Loewen, J., et al.: Dynamic CT Angiography in Vascular Imaging: Principles and Applications. RadioGraphics **42**(7), E224–E225 (2022)
19. Shahzad, R., Pennig, L., Goertz, L., et al.: Fully automated detection and segmentation of intracranial aneurysms in subarachnoid hemorrhage on CTA using deep learning. Scientific Reports **10**(1), 21799 (Dec 2020)
20. Shi, Z., Miao, C., Schoepf, U.J., et al.: A clinically applicable deep-learning model for detecting intracranial aneurysm in computed tomography angiography images. Nature Communications **11**(1), 6090 (Nov 2020)
21. Silversmith, W., Hilei, P.: 3D Euclidean Distance Transform. https://github.com/seung-lab/euclidean-distance-transform-3d (2019)
22. Song, T., Chen, J., Luo, X., et al.: CPM-Net: A 3D Center-Points Matching Network for Pulmonary Nodule Detection in CT Scans. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020, vol. 12266, pp. 550–559. Springer International Publishing, Cham (2020)
23. Terasaki, Y., Yokota, H., Tashiro, K., et al.: Multidimensional Deep Learning Reduces False-Positives in the Automated Detection of Cerebral Aneurysms on Time-Of-Flight Magnetic Resonance Angiography: A Multi-Center Study. Frontiers in Neurology **12**, 742126 (Jan 2022). https://doi.org/10.3389/fneur.2021.742126

24. Vaswani, A., Shazeer, N., Parmar, N., et al.: Attention is all you need. In: Advances in Neural Information Processing Systems. vol. 30 (2017)
25. Wang, J., Sun, J., Xu, J., et al.: Detection of Intracranial Aneurysms Using Multiphase CT Angiography with a Deep Learning Model. Academic Radiology **30**(11), 2477–2486 (Nov 2023). https://doi.org/10.1016/j.acra.2022.12.043
26. Wei, X., Jiang, J., Cao, W., et al.: Artificial intelligence assistance improves the accuracy and efficiency of intracranial aneurysm detection with CT angiography. European Journal of Radiology **149** (Apr 2022)
27. Wu, K., Gu, D., Qi, P., et al.: Evaluation of an automated intracranial aneurysm detection and rupture analysis approach using cascade detection and classification networks. Computerized Medical Imaging and Graphics **102**, 102126 (Dec 2022)
28. Xie, Y., Jiang, H., Gkioxari, G., Straub, J.: Pixel-Aligned Recurrent Queries for Multi-View 3D Object Detection. In: 2023 IEEE/CVF International Conference on Computer Vision (ICCV). pp. 18324–18334. IEEE, Paris, France (Oct 2023)
29. Yang, J., Xie, M., Hu, C., et al.: Deep Learning for Detecting Cerebral Aneurysms with CT Angiography. Radiology **298**(1), 155–163 (Jan 2021)
30. You, W., Sun, Y., Feng, J., et al.: Protocol and Preliminary Results of the Establishment of Intracranial Aneurysm Database for Artificial Intelligence Application Based on CTA Images. Frontiers in Neurology **13** (2022)
31. Zhu, G., Luo, X., Yang, T., et al.: Deep learning-based recognition and segmentation of intracranial aneurysms under small sample size. Frontiers in Physiology **13**, 1084202 (Dec 2022). https://doi.org/10.3389/fphys.2022.1084202
32. Zhu, X., Su, W., Lu, L., et al.: Deformable Transformers for End-to-End Object Detection (2021). https://doi.org/10.48550/arXiv.2010.04159