2024 9th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM) | 979-8-3315-0448-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/SEEDA-CECNSM63478.2024.00037

Aorta localization in Computed Tomography images: A YoloV9 segmentation approach

Panagiotis N. Smyrlis, Gabriela R. Stancu, Foteini Dimaraki, Aimilia Ntetska,
Thomai Karamitsou, Katerina D. Tzimourta, Markos G. Tsipouras, Pantelis Angelidis
Laboratory of Biomedical Technology and Digital Health
School of Electrical and Computer Engineering
University of Western Macedonia
Kozani, Greece
pan.smyrlis@gmail.com, gabstancu@outlook.com, fodim96@gmail.com, antetska@uowm.gr,
tkaramitsou@uowm.gr, ktzimourta@uowm.gr, mtsipouras@uowm.gr, paggelidis@uowm.gr

***Abstract*—In this study, state-of the art segmentation methods were tested for use in biomedical CTA images to precisely highlight aorta instances. A set of YOLO architectures was trained on a newly annotated dataset of 120 transverse view images and detailed cross-validation results were obtained to measure the detection and precise localization ability of the models. The recent YOLOV9 algorithm flavors V9-C and V9-E were studied and compared to the previous V8-X model. The V9 version did score remarkable results to Mean Average Precision metrics, achieving 76.15% segmentation mAP50-95 score in the YOLOV9-C testing.**

***Keywords*—Computed Tomography, Aorta, Convolutional Neural Network, YOLO-V9, Segmentation, Medical Imaging**

I. Introduction

The accurate localization of the aorta in Computed Tomography (CT) images is a critical task in medical imaging, serving as a fundamental step for various diagnostic and therapeutic procedures[1]. Precise identification and segmentation of the aorta may significantly enhance the detection and treatment of aortic diseases, such as aneurysms, dissections, and atherosclerosis, thereby improving patient outcomes, while recent advances in visual detection of aortic maligant tissue using CT can nowadays enhance the reliable diagnosis' accuracy.

Recent advances in sensor tecnology, e.g. Photon-Counting CT (PCCT) have now achieved significant reduction for the patients' radiation exposure. PCCT enables higher resolution imaging using multi-energy images, trying to provide clearer and richer screening on images, including the organs and tissue around aorta, thus presenting a potential to monitor cases like post-surgery complications of aortic aneurysm repair [2].

Recent applications in literature, try to highlight the aortic sarcoma on Computed Tomography images[3], using automated image processing and analysis methods to precisely delineate the cancer presence. Aortic sarcoma may be seen as an increasing number of cancer cells on the aorta walls, while Aortic angiosarcoma, a distinct maligant tumor is present in aorta vessels. The automation of detecting these rare cancers employs conventional or advanced image processing methods like segmentation, 3D reconstruction and contrast enhancement, to spot and quantify tumors of interest, while the CT imaging is used, as providing a descriptive data source on this task.

A silent and progressive aorta disease is the Abdominal Aortic Aneurysm (AAA). It is seen as a gradual dilatation and eventual rupture of the abdominal aorta. The AAA has a high mortality rate and increasing prevalence along aging. The enstaging involves transformation of smooth muscle cells into an aberrant phenotype, with features like secretory phenotype, elevated differentiation marker miR-145, senescence marker SIRT-1, multi-nucleated and aberrant nuclear morphology, and higher levels of DNA damage marker γH2AX, which are potential indicators for pathological premature vascular aging. Most AAAs are asymptomatic until rupture, at which point severe abdominal or back pain, hypotension and a pulsatile abdominal mass occur. Back pain may be an early symptom in some cases.

Regular screening is crucial to detect AAAs before rupture, as their growth may be rapid or slow / without symptoms. Diagnostic tests include CT/MR Angiography, Duplex Ultrasound, CT scans and clinical examination. Studies on CT Angiography have yet presented a 95% sensitivity, 90% specificity and 92% accuracy score in detecting significant stenosis. MR Angiography, and particularly 3D Gadolinium-enhanced, is also highly effective[4]

This study project aims to develop a Digital Aortic Twin using computational tools and machine learning models fed by medical data to generate non-invasive biomarkers and predictive models for AAA rupture risk and geometric evolution. This can improve AAA management through advanced technologies and data analysis.

Machine learning holds great promise for improving the diagnostic accuracy and efficiency of CT and MR imaging for Abdominal Aortic Aneurysm disease. Recent advances in object detection and segmentation have presented remarkable ability to precisely localize and delineate objects of interest, including applications in medical imaging[5] and Computed Tomography. The YOLO series algorithms, using Convolutional Neural Networks and data augmentation have yet attracted remarkable interest, due to achieving state-of-the-art scores at a low computational cost. The YOLO-V9[6] version,

979-8-3315-0448-9/24/$31.00 ©2024 IEEE
DOI 10.1109/SEEDA-CECNSM63478.2024.00037

has recently presented an optimized approach using a novel architecture on top of previous yolo codebases and managed to score higher in detection and segmentation tasks for common objects.

In this study, the recent YOLO architecture is tested to segment CT images and localize the aorta instances, using a small amount of data. The model performance is compared to the preceeding architecture and detailed evaluation metrics are resided.

II. Materials and Methods

A. Dataset

To form the dataset, a human Abdominal CT scan was obtained as a sequence of Tomography scanner input following the Digital Imaging and Communications in Medicine (DICOM) standard. Subsequently, the data were sampled as 3-channel RGB visual images and manually annotated pixel-wise. A detailed snapshot of the initial dataset, which contains all distinct views (transverse, sagittal, coronal) for a random image, together with the whole scan 3d view is resided in Figure 1.

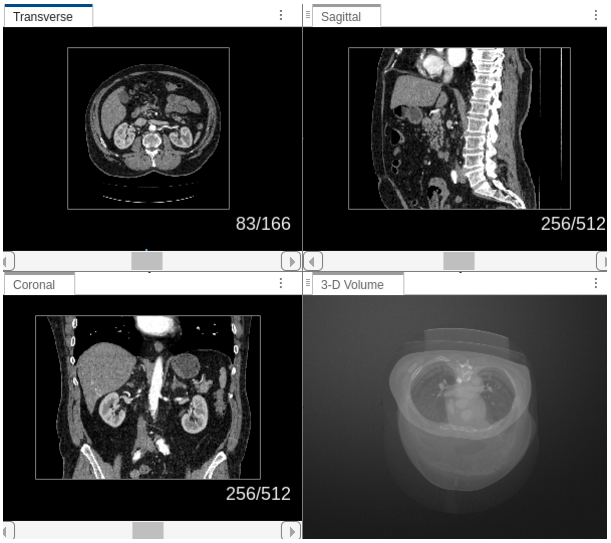


Fig. 1: The DICOM dataset view along each distinct axis and their 3d view

The dataset used for the experimental procedure contained 120 of the transverse view images from the above human CT scan. All images were manually annotated for the category of interest (aorta) and pixel-level masks were extracted. Figure 3 shows 4 random dataset images with their annotation overlaid in red color. All 120 annotated images contain one aorta instance and were fully annotated for the category of interest. The images are of size 512x512 pixels and the number of channels is 3, coded as RGB images.

B. The Yolov9 Model

The YoloV9[6] is the recent evolvement of Yolo series algorithm. In this version, its authors focus on better exploiting

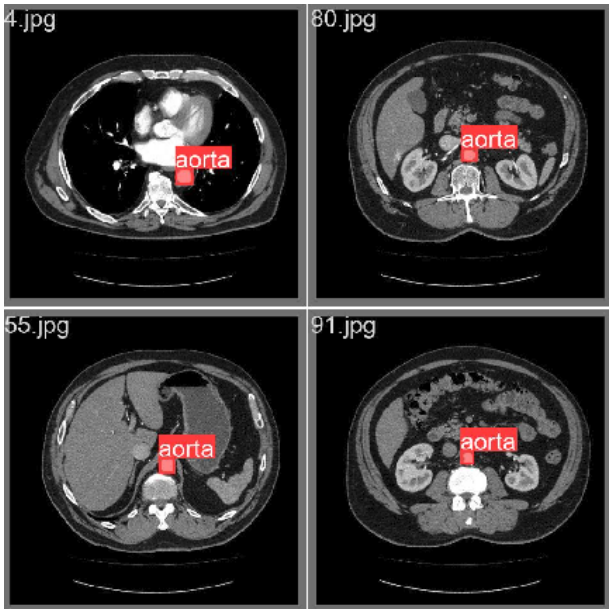


Fig. 2: Snapshots of the dataset and the annotation

the data information when raw data is modelled within the deep learning hidden layers. They examine better feature extraction methods to end up with the most representative model, and propose a number of novelties to Yolo architectures, starting from YoloV7[7] base concept. The v9 authors examine and embody novel architecture modules, Generalized ELAN and Programmable Gradient Information to formulate the algorithm. Both modules try to improve the detection ability and precision for objects of interest. They were extensively tested as standandalone and cooperative modules using common object data from the COCO dataset[8].

The PGI aims to mitigate the semantic loss when some data goes through multiple convolutional layers and optimize precise feature extraction. It uses a 3-branch schema, inspired by the reversible functions' principal, where some given function together with its inverse can transform data withought information loss. The module uses a main branch, which parameters are to be optimally learned by sharing into the following: 1. Auxiliary Reversible Branch and 2. Multi-level Auxiliary Information Branch. At last, the model inference is performed by the main branch only, while the other branches facilitate the accurate training procedure, resulting in a lower inference cost.

Subsequently, the authors extend the ELAN[9] architecture using CSPNet[10] modules. In result, the ELAN layer-stacking schema was further improved by using stronger computational blocks into GELAN. The network is designed to train with gradient path planning and was built extending the YoloV7 general schema, which was modified to facilitate anchor-free predictions. A diagram for the proposed network is shown in Figure 3.

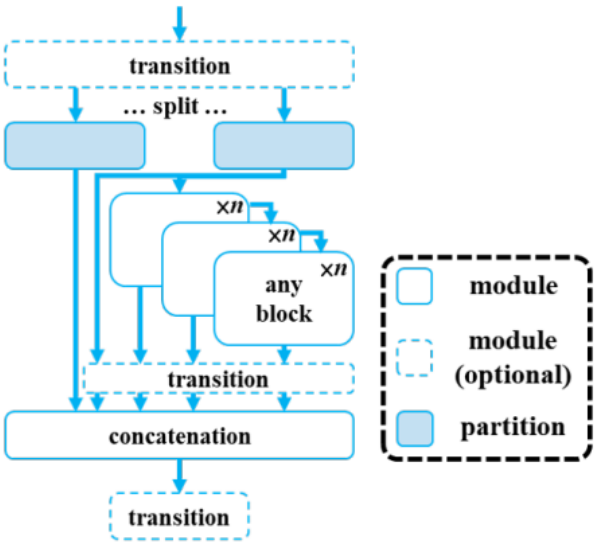


Fig. 3: GELAN architecture [6]

C. The experiment set up

The experimental procedure involved training two of YOLO V9 published architectures, and the YOLO V8-X architecture, which is considered as the top of V8 flavors. The learning strategy was identically set for all models, and the default model parameterization was kept as it was proposed by the Yolo authors. The image size was set to 512 pixels and the batch size was set to 8. The models were trained for 100 epochs.

K Fold Cross validation was employed with K=10, and the results presented are the cross validated metrices for the whole dataset (120 images). The testing set for each fold contained 12 images, while the remaining 108 were used in training.

III. Experimental Results

The metrices obtained were Precision, Recall, Mean Average Precision at threshold 50 and Mean Average Precision at 50-95 thresholds, measured both for the detection and segmentation algorithm performance.

In general, the V9-C flavor achieved to produce highest score in six out of 8 metrices, while followed by V8-X model which scored higher in Precision in both cases. The detailed metrices are shown in Table I. The V8 scored about 0.2% higher for segmentation and object detection Precision, while V9-C architecture achieved larger differences to the other metrices (up to 2.5% for mAP50-95). Both algorithms scored equally in Recall, where they manage to locate all instances required in this experiment. Moreover, the V9-C architecture proved to produce stronger decisions, as dictated by mAP50-95 metric, where it outperformed the V8-X model.

The metrices obtained show that the V9 architecture achieves higher performance scores and enhanced precision for segmentation and detection in the data tested. The V9-C model achieved to both to locate the category of interest and obtain stronger predictions. On the other hand, the increase of model size in V9-E flavor did not result in performance enhancement to the scope of this experiment.

TABLE I: Results obtained for all models and metrices

	Metric	YOLO V8-X	YOLO V9-C	YOLO V9-E
BOX	**Precision**	**99.19**	98.97	97.41
	Recall	**100.00**	**100.00**	99.96
	mAP50	99.49	**99.50**	99.20
	mAP50-95	82.75	**85.13**	79.43
MASK	**Precision**	**99.19**	98.97	97.41
	Recall	**100.00**	**100.00**	99.96
	mAP50	99.49	**99.50**	99.20
	mAP50-95	75.13	**76.15**	72.26

The models' training was gpu-accelerated with the use of an Nvidia RTX 3060 graphics card. Also, the model time performance was measured for all models and the execution time (in milliseconds) is shown in Table II, along the model size (number of parameters in millions). A visual comparison of model inference time towards the number of parameters is shown in Figure 4. The V9-C model, which scored best for the detection and segmentation tests, also achieved the best time and provided the most efficient solution in terms of computational cost.

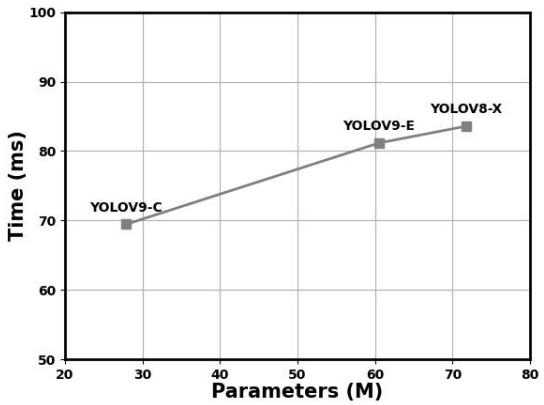


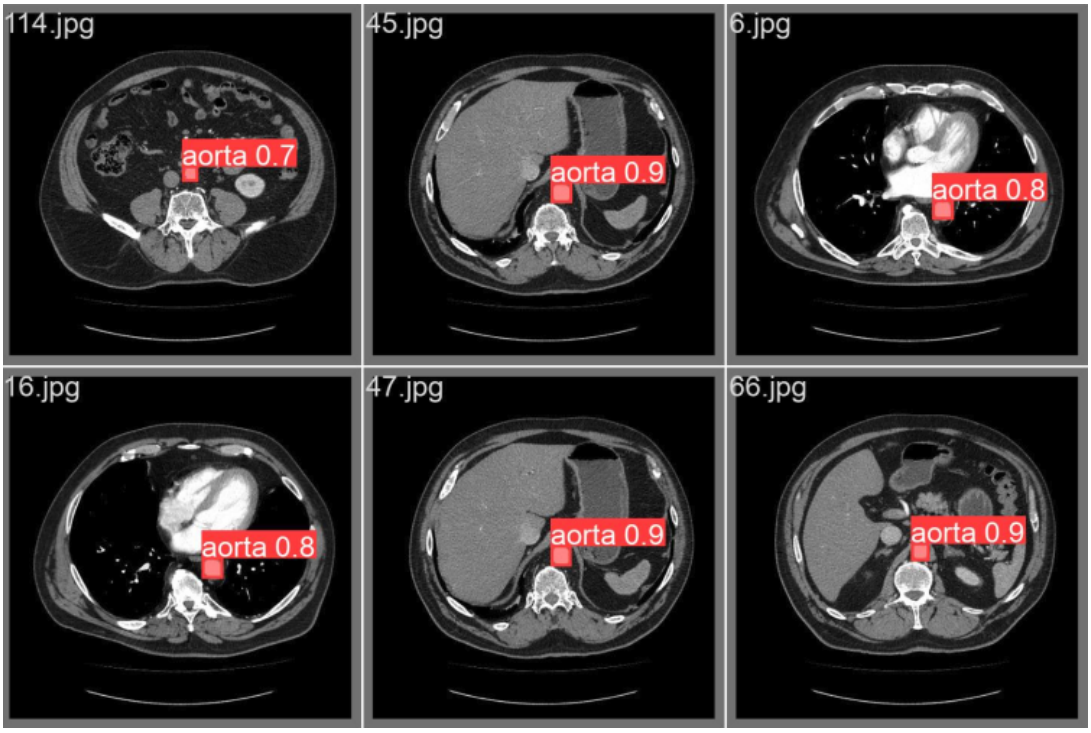
Fig. 4: Model inference time (ms per image) compared to the model parameters (million)

TABLE II: Model Parameters (Millions) and per image Inference Time (ms)

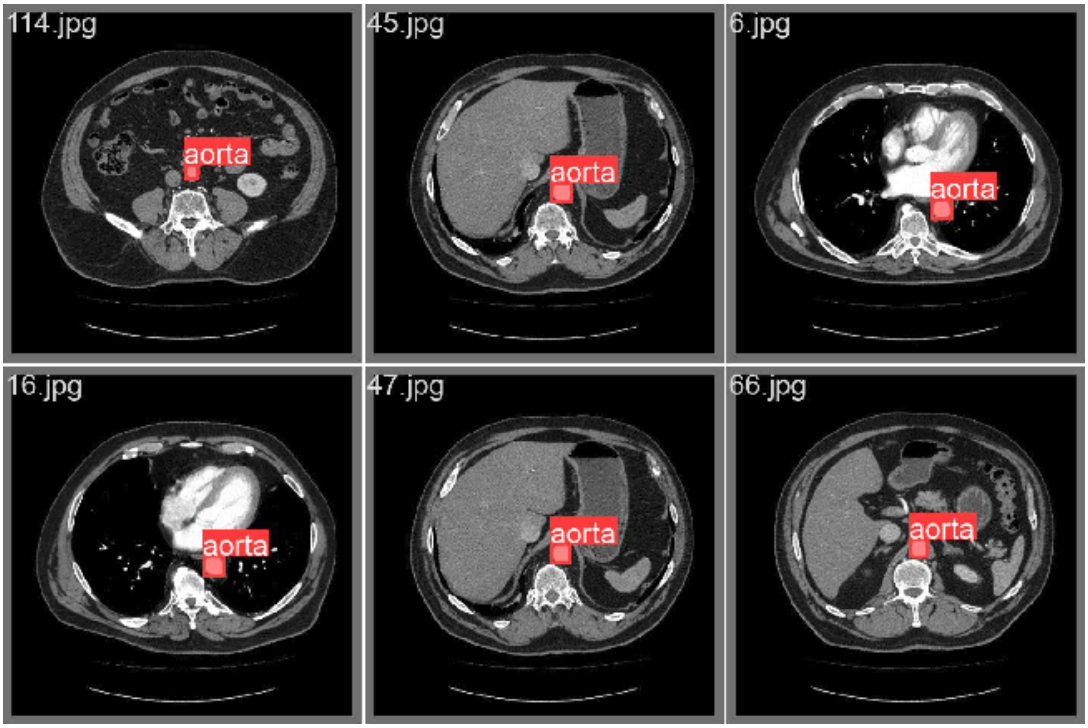
	YOLO V9-C	YOLO V9-E	YOLO V8-X
Parameters (M)	27.90	60.50	71.80
Time (ms)	69.45	81.16	83.58

IV. Conclusion

To provide a thorough assessment, we compared the previous Yolov8-X model with two state-of-the-art architectures from the YoloV9 model: YoloV9-C and YoloV9-E. We used a set of 120 CT biomedical images in the experimental



(a)



(b)

Fig. 5: Example of YOLO V9-C predictions to random CT images (a) and their Ground truth (b)

process to precisely highlight aorta instances. All models followed the same learning approach, maintaining the default parameterization set by the Yolo authors. Furthermore, the models underwent 100 epochs of training with 10-fold cross-validation. The metrics used to evaluate the models were Precision, Recall, Mean Average Precision at a 50 threshold (mAP50), and Mean Average Precision at 50-95 thresholds (mAP50-95), as shown in Table I.

In the tested data, the YoloV9 architecture performed better overall, demonstrating improved precision for segmentation and detection. The YoloV9-C model was particularly effective at locating the category of interest and achieving more accurate enhancement. In contrast, the larger YoloV9-E model did not show performance improvements. As seen in Fig 4, the YoloV9-C model also had the shortest inference time (milliseconds per image) relative to the number of model parameters.

Our study introduces a novel approach to accurately highlight aortic instances in CT images. We applied an effective instance segmentation architecture based on the YoloV9 design. This approach achieved high performance and improved precision for aorta segmentation and detection using the lightweight model. The results obtained with the recent Yolo architecture on this specific dataset are promising, demonstrating its applicability to biomedical image segmentation problems.

ACKNOWLEDGMENT

This research has been financed by the Operational Program "Flagship actions in interdisciplinary scientific fields with a special focus on the productive fabric", under the call Recovery & Resilience Facility - RRF (project code: TAEDR-0535983).

REFERENCES

[1] D. Andreini, E. Martuscelli, A. I. Guaricci, N. Carrabba, M. Magnoni, C. Tedeschi, A. Pelliccia, G. Pontone *et al.*, "Clinical recommendations on cardiac-ct in 2015: a position paper of the working group on cardiac-ct and nuclear cardiology of the italian society of cardiology," *Journal of Cardiovascular Medicine*, vol. 17, no. 2, pp. 73–84, 2016.

[2] A. Meloni, F. Frijia, D. Panetta, G. Degiorgi, C. De Gori, E. Maffei, A. Clemente, V. Positano, and F. Cademartiri, "Photon-counting computed tomography (pcct): technical background and cardio-vascular applications," *Diagnostics*, vol. 13, no. 4, p. 645, 2023.

[3] G. Wu, R. Xie, Y. Li, B. Hou, J. N. Morelli, and X. Li, "Histogram analysis with computed tomography angiography for discriminating soft tissue sarcoma from benign soft tissue tumor," *Medicine*, vol. 99, no. 2, p. e18742, 2020.

[4] H. Joshi, R. Shah, J. Prajapati, V. Bhangdiya, J. Shah, Y. Kandre, and K. Shah, "Diagnostic accuracy of computed tomography angiography as compared to conventional angiography in patients undergoing non-coronary cardiac surgery," *Heart Views*, vol. 17, no. 3, pp. 88–91, 2016.

[5] P. N. Smyrlis, O. Tsakai, K. Vogklis, N. Giannakeas, A. Tzallas, G. F. Fragulis, and M. G. Tsipouras, "A comparative testing of yolov8-based algorithm for cancer biopsy images: Case study on pannuke pan-cancer dataset," in *2023 8th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM)*. IEEE, 2023, pp. 1–5.

[6] C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, "Yolov9: Learning what you want to learn using programmable gradient information," *arXiv preprint arXiv:2402.13616*, 2024.

[7] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2023, pp. 7464–7475.

[8] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*. Springer, 2014, pp. 740–755.

[9] C.-Y. Wang, H.-Y. M. Liao, and I.-H. Yeh, "Designing network design strategies through gradient path analysis," *arXiv preprint arXiv:2211.04800*, 2022.

[10] C.-Y. Wang, H.-Y. M. Liao, Y.-H. Wu, P.-Y. Chen, J.-W. Hsieh, and I.-H. Yeh, "Cspnet: A new backbone that can enhance learning capability of cnn," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, 2020, pp. 390–391.