Anatomy-Aware MAE Pretraining for Generalizable Coronary Artery Segmentation

Miriam Gutiérrez-Fernández[1], Achilleas Toumpas[2], Christoniki Maga-Nteve[2], and Karen López-Linares Román[1,3]

[1] Vicomtech Foundation, Basque Research and Technology Alliance (BRTA), Mikeletegi 57, 20009 Donostia-San Sebastián, Spain
[2] Information Technologies Institute (ITI), Centre for Research and Technology Hellas (CERTH), 57001, Thessaloniki, Greece
[3] eHealth Group, Bioengineering Area, Biogipuzkoa Health Research Institute, San Sebastián, Spain

Abstract. Automated coronary artery segmentation from X-ray coronary angiography (XCA) remains limited by scarce and heterogeneous annotations. Building on anatomy-guided masked autoencoder (MAE) pretraining, we present a systematic ablation study across three axes: the degree of anatomy-guided masking, the pretraining objective (reconstruction only, reconstruction with a vascular topology-preservation loss, and a further feature-consistency objective), and the fine-tuning strategy (full, partial, frozen). The encoder is pretrained on over 65,000 unlabelled XCA frames and transferred to downstream segmentation using 1,200 labelled images from ARCADE. We evaluate in-domain performance on a held-out test set and out-of-domain performance on an external dataset (DCA1), showing that pretraining consistently improves segmentation metrics over supervised baseline. Notably, full loss pretraining with partial fine-tuning and low anatomy-guided masking achieved the strongest balanced in-domain performance, reaching mean Dice of 0.77 and mclDice of 0.85 on in-domain dataset. Cross-domain analysis showed fine-tuning depth benefited both domains, whereas increasing pretraining objective complexity helped only in-domain at low-to-moderate masking and degraded both domains at high masking. Overall, our results show that MAE-based pretraining improves coronary segmentation robustness, with reconstruction-only pretraining, moderate anatomy-guided masking, and full or partial fine-tuning consistently yielding the most robust transfer in and out of domain datasets.

Keywords: Coronary angiography · Coronary artery segmentation · Self-supervised learning · Pretraining · MAE

1 Introduction

X-ray coronary angiography (XCA) is the gold standard for coronary vasculature assessment, supporting stenosis quantification, Coronary Artery Disease diagnosis, and risk stratification [18]. XCA-derived measurements also hold broader

research potential as proxy markers linking coronary phenotype to poorly understood cardiovascular outcomes [23]. Despite this value, XCA analysis remains challenging: images are two-dimensional projections of complex three-dimensional vascular networks, prone to vessel overlap and acquisition variability. Manual quantification is thus time-consuming and operator-dependent [5].

Deep learning has established itself as the leading approach for automated coronary artery segmentation from XCA, producing the structured vascular representations needed for quantitative analysis at scale. Over the past decade, supervised methods have dominated the field, relying on either radiologist-annotated masks or pseudo-masks extracted semi-automatically via classical approaches such as the Frangi filter [10]. Most published work builds on well-established encoder-decoder architectures, with U-Net and its variants being the predominant choice [26, 20, 24, 2]. Beyond architectural design, topology-aware loss functions have been introduced to better preserve vascular structure during training. These include skeleton-based losses [25], connectivity-preserving losses [12], and focal loss variants adapted for tubular structures [14, 1].

However, supervised approaches depend heavily on high-quality annotations, which remain scarce. Labelling conventions vary substantially across public datasets, with some providing only coarse major-vessel delineations, while others offer dense pixel-level annotations. These inconsistent, heterogeneous annotations introduce contradictory supervision signals for the same anatomical structures, undermining model generalisation. As a result, there is a growing interest in self-supervised learning (SSL) to reduce the dependence on manual annotations. SSL broadly encompasses contrastive methods, which learn by pulling together augmented views of the same image while pushing apart views from different images [7, 8], and self-prediction methods such as Masked Autoencoders (MAE) [13], which learn by reconstructing randomly masked image regions.

Early evidence has demonstrated the potential of applying these paradigms to XCA through frame-level self-supervised pretraining. However, transferring SSL methods developed for natural images to this domain requires careful adaptation, as standard augmentation strategies and random masking can inadvertently destroy diagnostically relevant vascular structures, where fine vessel detail carries direct clinical value [16]. CM-UNet [6] uses contrastive-MAE pretraining on the 1,738-image FAME2 dataset. Fine-tuning on 18 versus 500 images caused only a 15.2% Dice score drop, compared to 46.5% for non-pretrained baselines. However, the small pretraining dataset limits generalizability conclusions. VasoMIM [15] introduced a MAE pretraining framework enhanced with anatomy-guided masking and an anatomical consistency loss, trained on 170k images across six datasets, achieving state-of-the-art transferability. However, its reliance on Frangi-based pseudo-labels provides only coarse structural guidance.

While anatomy-guided masking [15] and segmentation-aware pretraining losses [11] have each been explored independently, their interaction, and the extent to which each component contributes to downstream transferability, remains poorly understood. In this work, we present a systematic study of anatomy-aware MAE pretraining for coronary artery segmentation, in which we isolate

and evaluate three factors: *(i)* the degree of anatomy-guided masking, driven by pseudo-labels from a pretrained segmentation model [11]; *(ii)* the pretraining objective, comparing plain reconstruction, reconstruction augmented with a segmentation-consistency loss [11] targeting tubular vascular structure, and a further addition of a feature-consistency loss inspired by V-JEPA [3]; and *(iii)* the fine-tuning strategy, comparing full, partial, and frozen-encoder transfer. Pretraining is conducted on over 65,000 unlabelled angiographic frames, and the resulting encoders are fine-tuned on 1,200 labelled images. We evaluate downstream segmentation not only on an in-domain held-out test set, but also on an out-of-domain dataset, providing a rigorous assessment of generalisation.

2 Methods

2.1 Dataset

Table 1 summarises the datasets used for pretraining and downstream segmentation, spanning multiple vendors and preprocessed into a unified format retaining only contrast-visible frames.

Dataset	Role	Train	Val/Test	Vendor
CoronaryDominance [17]†	Pretrain	23,165	2,620	Philips Allura Clarity / Philips Azurion
DeepSA (LM-CAD) [27]	Pretrain	37,930	1,000	Unknown
XCAD [19]	Pretrain	1,459	162	GE Innova IGS 520
X-Ray-Syntax-Score [21]	Pretrain	3,028	405	Philips Allura Clarity / Siemens Axiom Artis
Pretrain Total		**65,582**	**4,187**	
ARCADE [22]	Downstream (train/val)	1,000	200	Philips Azurion 3 / Siemens Artis Zee
ARCADE [22]	Downstream (test)	–	300 *	
DCA1 [4]	Downstream (test)	–	134	Unknown

Table 1: Pretraining and downstream segmentation datasets. †Video sampled at k=5. *Same acquisition sites as ARCADE train/val.

2.2 Pretraining architecture and implementation

As pretraining **backbone**, we adopt ViT-B/16 [9] (patch size 16×16, D=768, 12 blocks, 12 heads, MLP ratio 4). Masked regions are zeroed in pixel space before patch embedding, so the encoder E always processes the full token sequence. A lightweight decoder (4 transformer blocks) receives the encoder feature map and reconstructs the full image via a linear projection head followed by unpatchify operation, as shown in Fig. 1A.

The reconstruction loss is inspired by the ViT-MAE masking-and-reconstruction strategy [13], adapted as dense pixel-space inpainting: masked patches are zeroed out before tokenization rather than dropped as tokens, so the encoder always sees a full-resolution input. We compare two masking strategies: random 32×32

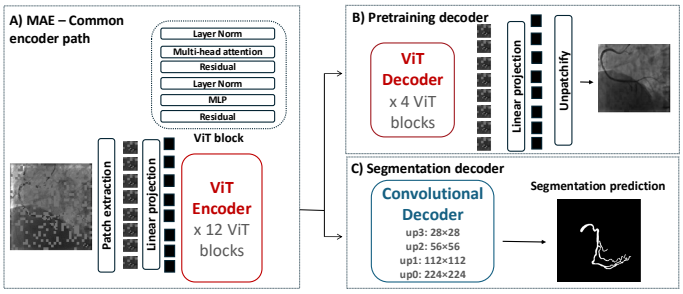


Fig. 1: Proposed architecture. **(A)** MAE ViT B/16 encoder for pretraining, **(B)** ViT decoder for pretraining, **(C)** Convolutional decoder for downstream segmentation.

patch masking with probability p_{rand}, and anatomy-guided 8×8 patch masking with probability p_{anat} restricted to patches overlapping vessel centerlines, extracted from a prior ARCADE-trained model [11]. Since p_{anat} is defined only over anatomically eligible patches, the two occlusion levels are not summative; the masks are combined as a pixel-wise union and may overlap.

Beyond masking, we assess how much additional anatomical and semantic guidance benefits pretraining by comparing three progressively richer loss objectives: reconstruction alone, reconstruction with a segmentation-consistency term, and reconstruction with both segmentation- and feature-consistency terms.

The **segmentation consistency loss** reuses FLOw-Loss, previously introduced for centerline-aware segmentation [11]: the reconstruction $\hat{x}$ is passed through a frozen segmentor S to obtain $\hat{y} = S(\hat{x})$, and the loss is computed between $\hat{y}$ and cached pseudo-labels, with gradients flowing back through S into the encoder to encourage anatomically consistent representations. We further implement an **in-house feature-consistency loss**, adapting JEPA's principle of predicting stable latent representations across masked and unmasked views: a target encoder E_{EMA} (an EMA copy of E) processes x and $\hat{x}$ into pooled feature maps z and $\hat{z}$, whose cosine similarity is maximised. Unlike V-JEPA's predictor-based objective, we use this simpler pooled-feature formulation to test whether representation-level supervision adds value beyond pixel- and segmentation-level signals.

The total pretraining loss, combining reconstruction, segmentation, and feature terms, can be expressed as $\mathcal{L} = \lambda_{rec}\mathcal{L}_{rec} + \lambda_{seg}\mathcal{L}_{seg} + \lambda_{feat}\mathcal{L}_{feat}$, where λ_{rec}, λ_{seg}, and λ_{feat} are set to 1.0, 0.1, and 0.1 respectively, chosen empirically based on early-epoch validation performance.

2.3 Segmentation architecture, implementation and evaluation

For downstream segmentation, the pretrained ViT encoder is coupled with a convolutional decoder using skip connections from blocks 2, 5, and 8, projected to 512, 256, 128, and 64 channels via 1×1 convolutions and fused through

transposed-convolution upsampling stages with double Conv-BN-ReLU blocks (Fig. 1C).

Three transfer strategies are evaluated against the same ViT-UNet trained from scratch (supervised baseline): *full fine-tuning* updates all encoder weights; *partial freezing* freezes early layers while updating deeper ones; and *full freezing* uses the encoder as a fixed feature extractor training only the decoder.

Segmentation is evaluated using Dice, clDice, their masked variants (mDice, mclDice) and precision, following [11]. Masked variants restrict evaluation to annotated regions, mitigating false-positive penalisation in partially annotated datasets. Statistical comparisons use the Wilcoxon signed-rank test ($p<0.05$).

2.4 Implementation details and ablation design

All pretraining experiments use the AdamW optimiser with a base learning rate of 1×10^{-4}, weight decay of 1×10^{-6}, and a cosine annealing schedule, for a total of 100 epochs with batch size 32. Input images are resized to 512×512; no additional data augmentation is applied beyond masking. The EMA decay for the target encoder is set to $\tau = 0.996$.

To isolate the contribution of each design choice, we ablate three factors independently. First, the **pretraining objective**: reconstruction only ($\mathcal{L}_{rec}$), reconstruction with segmentation consistency ($\mathcal{L}_{rec} + \mathcal{L}_{seg}$), and the full loss. Second, the **anatomy-guided masking proportion**, $p_{anat} \in \{0, 0.3, 0.5, 0.7\}$, with fixed $p_{rand} = 0.5$. Third, the **fine-tuning strategy** used to transfer the pretrained encoder to downstream segmentation: full fine-tuning, partial freezing, and full freezing of encoder weights. Pretraining was run on a single NVIDIA A40 GPU (48 GB), with 8 CPU workers for data loading.

For downstream segmentation, models are trained with AdamW ($lr = 1 \times 10^{-4}$, weight decay 1×10^{-6}) for 100 epochs, monitored on validation Dice, with batch size 32. Segmentation experiments were run on a single NVIDIA L40S GPU (48 GB), with 20 CPU workers and 56 GB RAM.

3 Results

Fig. 3 summarises segmentation performance across masking ratios, fine-tuning strategies, and pretraining loss configurations for Dice, clDice, mclDice and precision in the test datasets, with segmentation examples in Fig. 2.

Overall, pretraining outperformed the supervised ARCADE baseline (dotted line, Fig. 3) in the majority of the 72 configurations tested, with consistent gains in vessel recovery and topological completeness (Dice, clDice, mclDice), though improvements in Precision were less uniform and more sensitive to the specific loss, masking, and fine-tuning strategy. The best configuration varied by dataset: on ARCADE, the full loss with partial fine-tuning and low anatomy-guided masking (p_{anat}=0.3) gave the strongest balanced performance (Dice: 0.77; mclDice: 0.85), although reconstruction-only pretraining with partial fine-tuning and intermediate anatomy-guided masking ($p_{anat} = 0.5$) achieved the

best masked vessel performance (mDice: 0.830; mclDice: 0.868), as shown qualitatively in Fig. 2. On DCA1, reconstruction-only pretraining without anatomy-guided masking and partial fine-tuning performed best (Dice: 0.65; mclDice: 0.81), though at the cost of lower precision.

Loss configuration. Averaging across all datasets, masking ratios, and fine-tuning strategies, reconstruction-only pretraining (L_{rec}) provided the strongest mean performance (Dice: 0.67; clDice: 0.74). The best configuration within each loss type supported this ranking, with L_{rec} reaching the highest global average under partial fine-tuning and p_{anat}=0.5 (Dice: 0.70; clDice: 0.79). This ranking held at the dataset level, with L_{rec} achieving the highest mean Dice and clDice on ARCADE and DCA1 individually. Notably, this advantage is largely driven by frozen-encoder transfer: on ARCADE and DCA1, freezing the encoder after L_{rec}+L_{seg} or full-loss pretraining caused a marked performance drop as p_{anat} increased, often falling below baseline, whereas L_{rec}-pretrained encoders remained stable and above baseline across all masking levels under frozen transfer. This suggests that reconstruction-only pretraining yields representations that transfer well even without fine-tuning, while the auxiliary segmentation- and feature-consistency terms require full or partial adaptation to be effective.

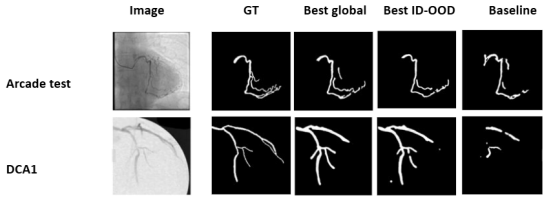


Fig. 2: Segmentation masks vs. ground truth (GT): supervised baseline, the best global configuration (L_{rec}, p_{anat}=0.5, partial fine-tuning) and the best external-domain configuration (L_{rec}, p_{anat}=0, partial fine-tuning).

Masking strategy. Overall, lower anatomy-guided masking ratios were more robust across domains. Averaging over all loss configurations, datasets, and fine-tuning strategies, p_{anat}=0 achieved the highest mean Dice (0.67) and clDice (0.74), followed monotonically by p_{anat}=0.3, 0.5, and 0.7 (Dice down to 0.60, clDice down to 0.68). This indicates that aggressive anatomy-guided masking over-constrains the representation, particularly under frozen-encoder transfer.

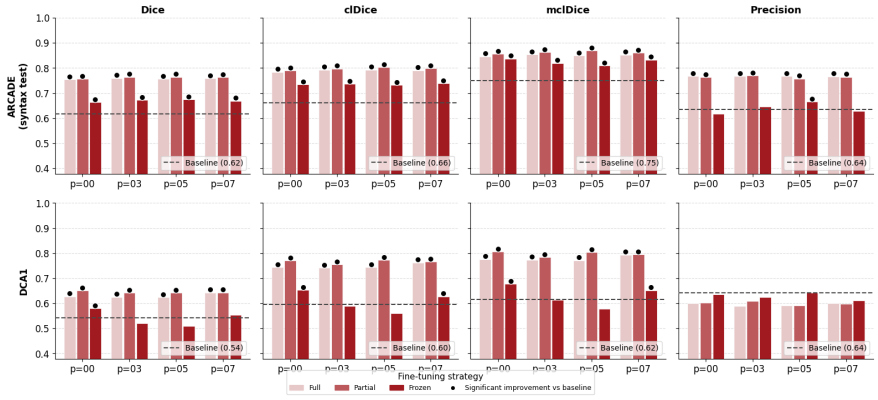
Fine-tuning strategy. Averaged across all loss configurations, masking ratios, and datasets, partial fine-tuning achieved the highest combined performance (Dice: 0.684, clDice: 0.755, Precision: 0.680), closely followed by full fine-tuning (Dice: 0.682, clDice: 0.755, Precision: 0.674), whereas frozen-encoder transfer was substantially lower (Dice: 0.551, clDice: 0.606, Precision: 0.655). Adapting the pretrained encoder consistently improved Dice, clDice, and Precision relative to frozen transfer on both datasets, with partial and full fine-tuning performing comparably throughout.

In-/out-of-domain consistency (ID-OOD). We evaluated cross-domain stability via the ID-OOD gap (Δ), defined as the absolute difference between ARCADE test and DCA1. Fine-tuning strategy had the largest effect: full and partial fine-tuning reached similar raw performance (Dice: 0.64–0.77 in-domain, 0.57–0.65 out-of-domain; clDice: 0.69–0.80 in-domain, 0.65–0.77 out-of-domain) and similar Δ, while the frozen backbone scored lower on both domains (Dice Δ up to 0.17). Pretraining objective determined how robust Δ was to masking. Reconstruction-only kept the Dice gap nearly flat across masking levels (Δ 0.12–0.13; clDice Δ 0.03–0.05). Adding segmentation and feature-consistency losses produced a wider gap at low-to-moderate masking (Dice Δ 0.14 at $p_{anat} \leq 0.3$) that narrowed sharply at $p_{anat}=0.7$ (Dice Δ 0.07), driven by ARCADE performance collapsing toward the DCA1 level rather than by improved transfer.

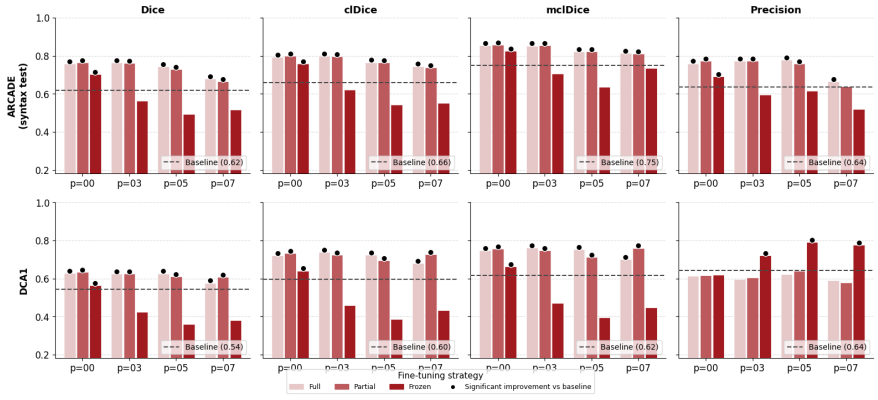
4 Conclusion

We presented a systematic study of anatomy-aware MAE pretraining for coronary artery segmentation in XCA, isolating the impacts of anatomy-guided masking, auxiliary objectives, and transfer strategies, extracting some key insights. First, **pretraining improved significantly** over the supervised baseline in most configurations. Second, **reconstruction-only MAE** gave the most robust transfer, suggesting reconstruction alone learns transferable vascular representations that benefit downstream adaptation. The segmentation-consistency term appears redundant with reconstruction, as the same segmentor generates both the pseudo-labels and the consistency target, while the feature-consistency term reflects an objective mismatch: comparing globally pooled representations collapses the spatial resolution a dense segmentation target requires, better suited to classification than to fine-grained vessel localisation. Third, lower-to-intermediate anatomy-guided masking best balanced in-domain performance and OOD consistency, consistent with VasoMIM [15], suggesting aggressive anatomical masking over-constrains the vessel representation. Fourth, adapting the pretrained encoder remained important: initialization alone was not enough without task-specific fine-tuning. Even if our work provides valuable insights into anatomy-aware MAE pretraining for coronary artery segmentation, several aspects should be considered when interpreting the results. The anatomical supervision may reflect biases from ARCADE-style annotations, since the frozen segmentor used to generate pseudo-labels was trained on that dataset. Future work could mitigate this effect by incorporating more diverse pseudo-label sources, for example by combining model-based labels with vessel-enhancement filters. Additionally, evaluation relies on a single external test set, although results may vary on datasets with different annotation patterns.

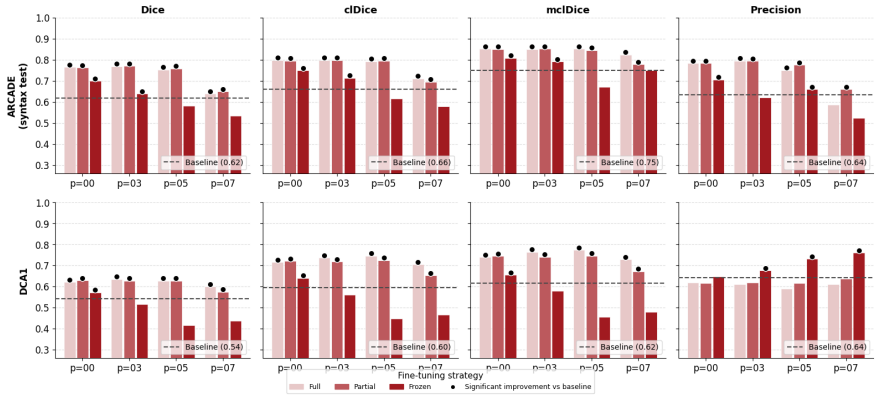
Acknowledgments. This research was funded by the EU's Horizon Programme (Grant N 101137278 – CVDLINK). The views expressed are solely those of the authors and do not represent the EU or the European Health and Digital Executive Agency, which are not responsible for them.



(a) Reconstruction only (L_{rec})



(b) Reconstruction + segmentation consistency ($L_{rec} + L_{seg}$)



(c) Full loss ($L_{rec} + L_{seg} + L_{feat}$)

Fig. 3: Segmentation performance by masking ratio and fine-tuning strategy, under three pretraining loss configurations. Dots indicate a significant improvement over the supervised baseline.

References

1. Acebes, C., Moustafa, A.H., Camara, O., Galdran, A.: The centerline-cross entropy loss for vessel-like structure segmentation: Better topology consistency without sacrificing accuracy. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 710–720. Springer (2024)
2. Algarni, M., Al-Rezqi, A., Saeed, F., Alsaeedi, A., Ghabban, F.: Multi-constraints based deep learning model for automated segmentation and diagnosis of coronary artery disease in x-ray angiographic images. PeerJ Computer Science **8**, e993 (2022)
3. Assran, M., Duval, Q., Misra, I., Bojanowski, P., Vincent, P., Rabbat, M., LeCun, Y., Ballas, N.: Self-supervised learning from images with a joint-embedding predictive architecture. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 15619–15629 (2023)
4. Cervantes-Sanchez, F., Cruz-Aceves, I., Hernandez-Aguirre, A., Hernandez-Gonzalez, M.A., Solorio-Meza, S.E.: Automatic segmentation of coronary arteries in x-ray angiograms using multiscale analysis and artificial neural networks. Applied Sciences **9**(24), 5507 (2019)
5. Chakrabarti, A.K., Grau-Sepulveda, M.V., O'Brien, S., Abueg, C., Ponirakis, A., Delong, E., Peterson, E., Klein, L.W., Garratt, K.N., Weintraub, W.S., et al.: Angiographic validation of the american college of cardiology foundation–the society of thoracic surgeons collaboration on the comparative effectiveness of revascularization strategies study. Circulation: Cardiovascular Interventions **7**(1), 11–18 (2014)
6. Challier, C., Sun, X., Mahendiran, T., Senouf, O., De Bruyne, B., Auberson, D., Müller, O., Fournier, S., Frossard, P., Abbé, E., et al.: Cm-unet: A self-supervised learning-based model for coronary artery segmentation in x-ray angiography. In: Annu. Int. Conf. IEEE Eng. Med. Biol. - Proc. pp. 1–7. IEEE (2025)
7. Chen, T., Kornblith, S., Norouzi, M., Hinton, G.: A simple framework for contrastive learning of visual representations. In: International conference on machine learning. pp. 1597–1607. PmLR (2020)
8. Chen, X., Xie, S., He, K.: An empirical study of training self-supervised vision transformers. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 9640–9649 (2021)
9. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al.: An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929 (2020)
10. Gao, Z., Wang, L., Soroushmehr, R., Wood, A., Gryak, J., Nallamothu, B., Najarian, K.: Vessel segmentation for x-ray coronary angiography using ensemble methods with deep learning and filter-based features. BMC Medical Imaging **22**(1), 10 (2022)
11. Gutiérrez-Fernández, M., Pérez-Herrera, L.V., Arrarte Terreros, N., López-Linares Román, K.: Flow-loss: A hybrid loss for centerline-aware segmentation in xca. In: AMAI 2025. Springer (2025)
12. He, H., Banerjee, A., Beetz, M., Choudhury, R.P., Grau, V.: Semi-supervised coronary vessels segmentation from invasive coronary angiography with connectivity-preserving loss function. In: 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI). pp. 1–5. IEEE (2022)
13. He, K., Chen, X., Xie, S., Li, Y., Dollár, P., Girshick, R.: Masked autoencoders are scalable vision learners. In: Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR). pp. 16000–16009 (2022)

14. Hu, K., Zhang, Z., Niu, X., Zhang, Y., Cao, C., Xiao, F., Gao, X.: Retinal vessel segmentation of color fundus images using multiscale convolutional neural network with an improved cross-entropy loss function. Neurocomputing **309**, 179–191 (2018)
15. Huang, D.X., Zhou, X.H., Gui, M.J., Xie, X.L., Liu, S.Q., Wang, S.Y., Xiang, T.Y., Ma, R.Z., Xiao, N.F., Hou, Z.G.: Vasomim: Vascular anatomy-aware masked image modeling for vessel segmentation. In: Proc. AAAI Conf. Artif. Intell. vol. 40, pp. 4985–4993 (2026)
16. Huang, S.C., Pareek, A., Jensen, M., Lungren, M.P., Yeung, S., Chaudhari, A.S.: Self-supervised learning for medical image classification: a systematic review and implementation guidelines. NPJ Digital Medicine **6**(1), 74 (2023)
17. Kruzhilov, I., Ikryannikov, E., Shadrin, A., Utegenov, R., Zubkova, G., Bessonov, I.: Neural network-based coronary dominance classification of rca angiograms. In: Doklady Mathematics. vol. 110, pp. S212–S222. Springer (2024)
18. Lawton, J.S., Tamis-Holland, J.E., Bangalore, S., et al.: 2021 ACC/AHA/SCAI guideline for coronary artery revascularization: A report of the american college of cardiology/american heart association joint committee on clinical practice guidelines. Circulation **145**(3), e18–e114 (Jan 2022)
19. Ma, Y., Hua, Y., Deng, H., Song, T., Wang, H., Xue, Z., Cao, H., Ma, R., Guan, H.: Self-supervised vessel segmentation via adversarial learning. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 7516–7525 (2021). https://doi.org/10.1109/ICCV48922.2021.00744
20. Mahendiran, T., Thanou, D., Senouf, O., Jamaa, Y., Fournier, S., De Bruyne, B., Abbé, E., Muller, O., Andò, E.: Angiopy segmentation: An open-source, user-guided deep learning tool for coronary artery segmentation. International journal of cardiology **418**, 132598 (2025)
21. Mahmoudi, S.S., Alishani, M.M., Emdadi, M., Hosseiniyan Khatibi, S.M., Khodaei, B., Ghaffari, A., Oskui, S.D., Ghaffari, S., Pirmoradi, S.: X-ray coronary angiogram images and syntax score to develop machine-learning algorithms for chd diagnosis. Scientific Data **12**(1), 471 (2025)
22. Popov, M., Amanturdieva, A., Zhaksylyk, N., Alkanov, A., Saniyazbekov, A., Aimyshev, T., Ismailov, E., Bulegenov, A., Kolesnikov, A., Kulanbayeva, A., et al.: Arcade: Automatic region-based coronary artery disease diagnostics using x-ray angiography images dataset (2023)
23. Sedoud, B., Caullery, B., Aouati, M., Marliere, S., Vautrin, E., Piliero, N., Ormezzano, O., Bouvaist, H., Lantuejoul, L.R., Vanzetto, G., et al.: Diagnostic relevance of angiography-derived coronary microcirculatory resistance in sudden cardiac death. Quantitative Imaging in Medicine and Surgery **16**(3), 234 (2026)
24. Shen, Y., Chen, Z., Tong, J., Jiang, N., Ning, Y.: Dbcu-net: deep learning approach for segmentation of coronary angiography images. The International Journal of Cardiovascular Imaging **39**(8), 1571–1579 (2023)
25. Shi, P., Hu, J., Yang, Y., Gao, Z., Liu, W., Ma, T.: Centerline boundary dice loss for vascular segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 46–56. Springer (2024)
26. Yang, S., Kweon, J., Roh, J.H., Lee, J.H., Kang, H., Park, L.J., Kim, D.J., Yang, H., Hur, J., Kang, D.Y., et al.: Deep learning segmentation of major vessels in x-ray coronary angiography. Scientific reports **9**(1), 16897 (2019)
27. Zeng, Y., Liu, H., Hu, J., Zhao, Z., She, Q.: Pretrained subtraction and segmentation model for coronary angiograms. Scientific Reports **14**(1), 19888 (2024)