

Automatic LV Localization and Short-Axis Plane Estimation from Arbitrary CMR Slice

Yi Yu¹, Yixuan Liu¹, Ziyu Zhang², Parker Martin¹, Zhenyu Bu¹,
Yuchi Han¹, and Yuan Xue¹(✉)

¹ The Ohio State University, Columbus, USA
Yuan.Xue@osumc.edu

² Nanjing University, Nanjing, China

Abstract. Accurate estimation of left ventricular (LV) orientation is essential for cardiac magnetic resonance (CMR) imaging and downstream analysis. Existing methods typically formulate orientation recognition as discrete view classification or rely on multi-slice geometric intersection, limiting their ability to model continuous 3D orientation and generalize across arbitrary slices. This work introduces a novel paradigm: Joint LV localization and 3D orientation estimation from a single CMR slice. To investigate this setting, representative orientation-aware detection frameworks are adapted to the CMR domain, and their limitations are analyzed. Upon that, we propose the Polar-Coupled Circular (PCC) embedding that provides a continuous and unambiguous orientation representation to address the limitations. Meanwhile, a scalable benchmark is constructed through automatic slice sampling from volumetric CMR segmentation datasets. Extensive experiments on four datasets demonstrate strong performance, achieving an average mIoU of 86.18% and an average angle deviation of 3.39°. This study establishes a new task setting for single-slice LV orientation modeling and provides a geometry-consistent framework for spatially informed CMR analysis. Code is available at <https://github.com/yuyi1005/cmr-3d-ood>.

Keywords: Cardiovascular magnetic resonance · Orientation estimation · Left ventricle localization.

1 Introduction

Cardiovascular magnetic resonance (CMR) plays a central role in quantitative cardiac assessment, providing high-resolution anatomical and functional information across multiple imaging planes [26]. In particular, evaluation of the left ventricle (LV) is fundamental for measuring cardiac function and diagnosing disease. Over the past decade, LV segmentation and quantification have been extensively studied and achieved substantial progress. These tasks highly rely on proper cardiac orientation, especially short-axis (SAX) orientation. Consequently, accurate modeling of **3D SAX orientation** is increasingly important for spatially aware analysis [2, 11], supporting applications such as cross-view registration, cardiac modeling, and automated plane prescription [5].

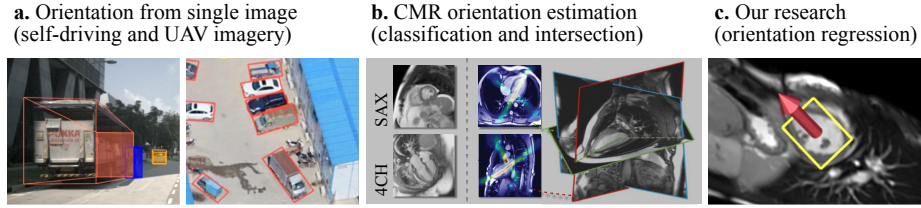


Fig. 1. Orientation recognition from a single image has achieved remarkable success in autonomous driving. However, in the CMR domain, orientation estimation still largely relies on traditional view classification or plane-intersection strategies.

Existing research on SAX orientation recognition, however, largely follows two paradigms: **1)** Some studies formulate it as a **view classification** problem, assigning images to predefined categories such as SAX, four-chamber (4CH), or two-chamber (2CH) views [1, 20, 7]. While effective for coarse categorization, they are inherently limited to discrete classes and cannot model continuous 3D orientation. **2)** Another line of work focuses on automated plane prescription, which estimates the cardiac orientation during acquisition [9]. They estimate the target plane through **geometric intersection** across multiple slices (see Fig. 1b) [15, 4, 8, 22]. Such approaches require multi-view aggregation.

Recent advances in autonomous driving have demonstrated that directly regressing 3D orientation from 2D observations is both feasible and highly effective. For instance, FCOS3D [21] can reliably localize vehicles and infer their 3D orientation from monocular images (see Fig. 1a). In CMR, despite the absence of perspective projection, different slice orientations likewise exhibit distinct anatomical patterns. This raises a natural question: Can 3D oriented object detection be extended to CMR? In this work, we explore this new direction and demonstrate that accurate LV localization and 3D orientation estimation from a single slice are achievable with strong performance (see Fig. 1c).

Our contributions are summarized as: **1) Task setting and benchmark.** We introduce a new problem setting: LV localization and 3D orientation estimation from a single CMR slice, and develop a data curation pipeline to construct the benchmark from CMR segmentation datasets. **2) Analysis of potential solutions.** We adapt four representative orientation-aware detection frameworks to the CMR setting, establishing strong baselines. Their behaviors are systematically analyzed to identify limitations in 3D SAX orientation estimation. **3) Methodological contribution.** To address the limitations of existing methods, we propose Polar-Coupled Circular (PCC) embedding that enforces unit-energy consistency, improving stability and accuracy in 3D orientation estimation.

2 Analysis of Applicable Methods

3D oriented object detection is an active research area, particularly in autonomous driving, where models estimate object orientation and spatial extent directly

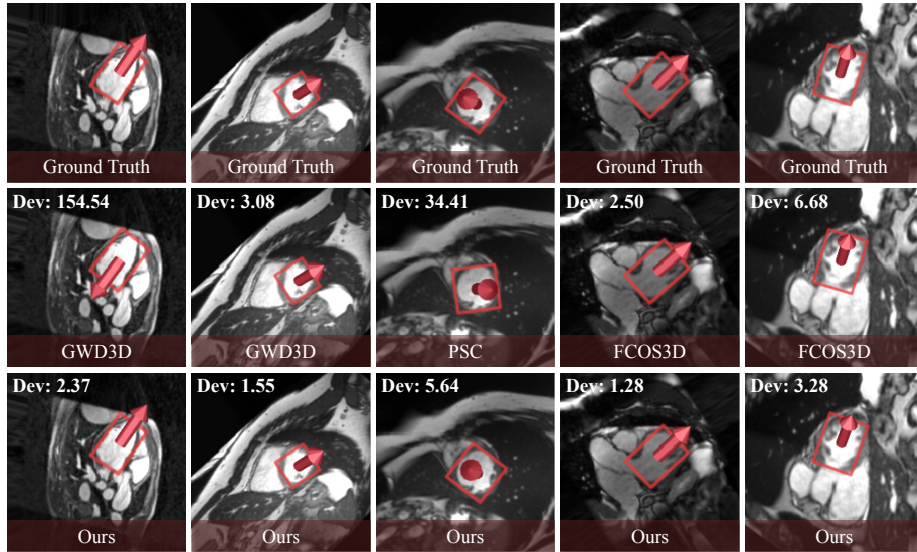


Fig. 2. Visualization of our work. Our model localizes the left ventricle with a rotated bounding box and estimates its 3D orientation (red arrows) using a single, arbitrary 2D slice. The second row displays some compared methods repurposed to our task.

from images [16]. Representative methods include: **1) L1-based.** Early methods directly regress 3D box parameters, including rotation angles, using L1 or smooth L1 losses. For example, FCOS3D [21] predicts orientation explicitly in parameter space, treating angle estimation as standard regression. **2) IoU-based.** Intersection over Union (IoU) optimizes geometric overlap instead of independent parameters. RIoU3D [28] introduces differentiable rotated IoU to supervise alignment directly in 3D geometric space. **3) Gaussian-based.** Gaussian-based methods reformulate oriented boxes as 2D/3D Gaussian distributions. GWD [23] measures similarity using Gaussian Wasserstein Distance, implicitly encoding orientation in the covariance matrix. **4) Encoding-based.** They encode oriented boxes into and regress encoded embeddings. PSC [24, 25] proposes boundary-free encoding to improve rotation stability during optimization.

Limitation analysis. The applicability of existing orientation-aware detection methods to CMR has not been previously studied. We therefore adapt representative approaches to our task. For methods originally designed for 2D rotation (e.g., PSC), we extend them to 3D by adding an additional regression channel for the z component of the normal vector, followed by a sigmoid activation to constrain the prediction to $[0, 1]$. As shown in Fig. 2, several limitations arise: **1) Ambiguity in IoU and Gaussian.** These methods are invariant to 180° rotation, making opposite directions indistinguishable. As a result, they cannot represent the full semi-spherical orientation space, leading to reversed predictions in certain cases. **2) Boundary issue in L1-based regression.** Due to angular periodicity, 180° and -180° are equivalent but yield a large L1 loss. This

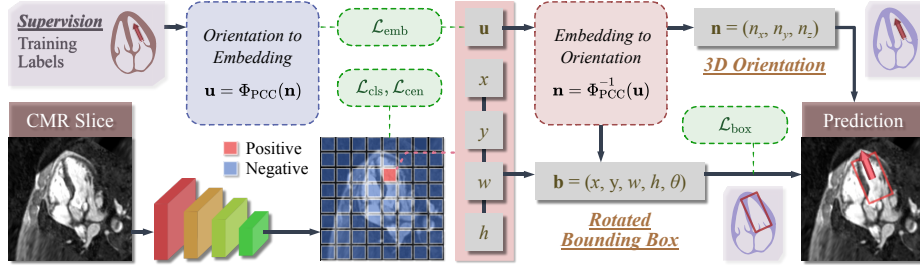


Fig. 3. The framework of our proposed model.

artificial discontinuity causes unstable optimization and degraded accuracy near boundaries. **3) Instability of PSC embedding.** Embedding-based methods such as PSC alleviate 2D boundary problem but are not inherently designed for 3D orientation. Extending them with an additional z -channel leads to varying embedding magnitudes across different directions and instability near polar regions (e.g., $z \rightarrow 1$), resulting in unreliable predictions.

Overall, existing formulations work well in LV localization, but exhibit representation ambiguity, boundary discontinuity, or instability when extended to continuous 3D SAX orientation regression in CMR.

3 Methodology

By analyzing the failure cases of existing methods (see Sec. 2), an ideal orientation model should satisfy the following properties: **1) Full coverage.** Its orientation spans the entire semi-sphere. **2) Uniform magnitude.** The embedding $\mathbf{n}$ satisfies $\|\mathbf{n}\|_2 = 1$ for all orientations. **3) Continuity.** No discontinuities across angular boundaries. **4) Uniqueness.** One-to-one correspondence between embedding and physical orientation, avoiding ambiguity.

In this section, we present our modified encoding-based method with all these features (see Fig. 3). The network adopts a ResNet-50 [10] with FPN [12] backbone, following the FCOS framework [19]. In addition to the standard bounding box parameters (x, y, w, h) , we introduce a Polar-Coupled Circular (PCC) embedding $\mathbf{u}$ for orientation prediction. During training, the embedding is computed from the ground-truth 3D normal vector as $\mathbf{u} = \Phi_{\text{PCC}}(\mathbf{n})$, as detailed in Sec. 3.1. The network is supervised to regress $\mathbf{u}$. During inference, the predicted embedding is decoded to recover the orientation $\mathbf{n} = \Phi_{\text{PCC}}^{-1}(\mathbf{u})$, as described in Sec. 3.2. The in-plane angle is then computed as $\theta = \arctan 2(n_y, n_x)$, and combined with (x, y, w, h) to form the rotated bounding box (x, y, w, h, θ) .

3.1 Polar-Coupled Circular (PCC) Embeddings

To obtain a continuous and geometrically consistent representation of 3D orientation, we introduce an PCC embedding, denoted as $\mathbf{u} = \Phi_{\text{PCC}}(\mathbf{n})$.

Given a unit normal vector $\mathbf{n} = (n_x, n_y, n_z)$, we encode it using evenly spaced circular harmonic bases. Let N be the encoding length. For $N \geq 4$, we compute:

$$p_n = \sqrt{\frac{2}{N-1}} (1 - n_z^2) \cos\left(\theta + \frac{2\pi n}{N-1}\right), \quad n = 0, \dots, N-2 \quad (1)$$

where $\theta = \arctan 2(n_y, n_x)$ and the last term $p_{N-1} = n_z$. The final embedding is constructed as:

$$\mathbf{u} = [p_0, \dots, p_{N-1}]^\top \quad (2)$$

For the special case of $N = 3$, the embedding is directly defined as $\mathbf{u} = (n_x, n_y, n_z)$, corresponding to the Cartesian representation of the normal vector. The scaling factor in Eq. (1) couples the azimuthal components with the polar coordinate, making the embedding satisfy the unit-energy constraint $\|\mathbf{u}\|_2^2 = 1$, ensuring that $\mathbf{u}$ lies on the unit hypersphere. This polar coupling suppresses azimuthal instability as $z \rightarrow 1$, where orientation about the polar axis becomes ambiguous, and provides a smooth and geometrically consistent representation for stable 3D orientation regression.

3.2 Recover Orientation from PCC Embeddings

Given a PCC embedding $\mathbf{u} = [p_0, \dots, p_{N-1}]^\top$, the reverse process $\mathbf{n} = \Phi_{\text{PCC}}^{-1}(\mathbf{u})$ recovers the 3D orientation vector $\mathbf{n} = (n_x, n_y, n_z)$ by projecting the circular harmonic components onto the first-order cosine and sine bases. Let N denote the encoding length. For $N \geq 4$, we compute the orthogonal projection:

$$(n_x, n_y) = \sqrt{\frac{2}{N-1}} \sum_{n=0}^{N-2} p_n \left(\cos \frac{2\pi n}{N-1}, -\sin \frac{2\pi n}{N-1} \right) \quad (3)$$

The polar component is directly preserved as $n_z = p_{N-1}$. The resulting vector $\mathbf{n} = (n_x, n_y, n_z)$ recovers the Cartesian representation of the 3D orientation.

3.3 Loss Functions

Based on FCOS [19], our method inherited a classification loss (Focal loss [13]), a box regression loss (IoU loss), and a center-ness loss (Binary Cross-Entropy loss). When PCC is incorporated, it involves an embedding loss. In the following, we introduce its calculation. First, the ground truth normal is encoded as:

$$\mathbf{u}_{\text{GT}} = \Phi_{\text{PCC}}(\mathbf{n}_{\text{GT}}) \quad (4)$$

where $\mathbf{u}_{\text{GT}}$ is the ground truth embeddings.

Afterward, the embedding loss can be calculated with L_1 loss:

$$\mathcal{L}_{\text{emb}} = L_1(\mathbf{u}_{\text{pred}}, \mathbf{u}_{\text{GT}}) \quad (5)$$

where $\mathbf{u}_{\text{pred}}$ is the output embeddings of the network.

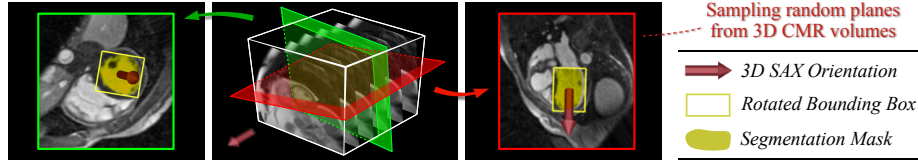


Fig. 4. Illustration of the data generation pipeline. Readily available 3D segmentation datasets are repurposed for our task setting, LV localization and oriented detection.

Finally, with the weighted losses $w_{\text{cls}}\mathcal{L}_{\text{cls}}$, $w_{\text{box}}\mathcal{L}_{\text{box}}$, and $w_{\text{cen}}\mathcal{L}_{\text{cen}}$ defined by backbone FCOS detector, the overall loss for training can be expressed as:

$$\mathcal{L} = w_{\text{cls}}\mathcal{L}_{\text{cls}} + w_{\text{box}}\mathcal{L}_{\text{box}} + w_{\text{cen}}\mathcal{L}_{\text{cen}} + w_{\text{emb}}\mathcal{L}_{\text{emb}} \quad (6)$$

where w_{emb} is set to 2 according to our ablation study (see Sec. 4).

3.4 Scalable Training Data Construction

For a newly defined task, the first and most essential contribution is the benchmark establishment. However, manually annotating 3D orientation from 2D images is extremely challenging. To address this limitation, we develop a data processing pipeline that automatically constructs a benchmark from volumetric CMR segmentation datasets, which are more widely available.

Given a 3D CMR volume with LV segmentation mask and the SAX direction (which is known for SAX stacks and manually annotated otherwise), we construct training samples through stochastic slicing (see Fig. 4). Specifically, we first randomly select a point P inside the LV mask. A normal vector $\mathbf{v}$, uniformly sampled from the unit sphere, is then used to define a plane Π passing through P . Since P lies within the LV mask, plane Π is guaranteed to intersect the LV, ensuring visible cardiac structures in the resulting slice. The volume is resampled onto plane Π to obtain a 256×256 2D CMR image, with the in-plane spatial resolution uniformly sampled between 1 mm and 2 mm. The original 3D SAX direction is then transformed into the coordinate system of this plane, yielding 3D orientation labels. For localization supervision, the minimum rotated bounding box enclosing LV mask is computed. The rotation of the box is aligned with the projection of the SAX direction onto plane Π .

By repeating this procedure, numerous randomly oriented 2D slices can be sampled from a single 3D volume. Given the availability of large-scale public segmentation datasets, this strategy is naturally scalable. In our experiments, 20 slices are sampled from each mask-annotated 3D CMR volume.

4 Experiments

Experiments are carried out on NVIDIA A100 GPUs using PyTorch 2.4.0 [18] and the rotation detection toolkit MMRotate 1.0.0 [29]. All experiments follow

Table 1. Comparison between our method and existing approaches on CMR datasets. Orientation accuracy is measured by angle deviation (in degrees, reported as mean $\pm$ std), localization accuracy by mIoU, and detection performance by AP₅₀ and AP₇₅, evaluated on the ACDC, M&Ms-2, In-House, and V2S-Real datasets.

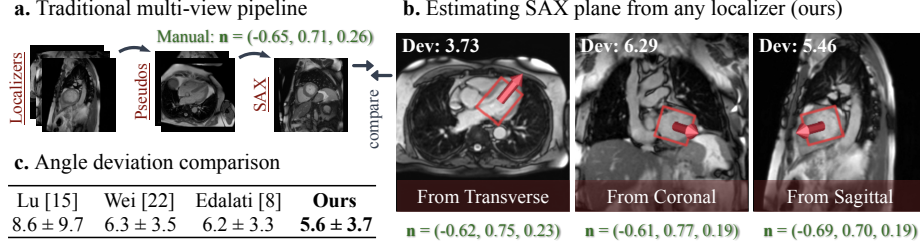
Methods	Angle Dev.	mIoU	AP ₅₀	AP ₇₅
▼ ACDC Dataset				
RIoU3D (2020) [28]	57.84 $\pm$ 46.51	82.94	99.27	74.17
FCOS3D (2021) [21]	4.87 $\pm$ 3.33	83.57	99.10	77.63
GWD3D (2023) [23]	58.73 $\pm$ 45.19	83.09	99.23	74.38
PSC (2024) [25]	4.49 $\pm$ 3.50	83.15	99.10	78.55
Ours	3.73 $\pm$ 2.82	83.96	99.11	80.47
▼ M&Ms-2 Dataset				
RIoU3D (2020) [28]	51.33 $\pm$ 40.77	83.66	99.89	78.18
FCOS3D (2021) [21]	4.94 $\pm$ 3.84	84.12	99.78	81.44
GWD3D (2023) [23]	52.26 $\pm$ 39.20	83.59	99.92	76.65
PSC (2024) [25]	4.67 $\pm$ 3.10	84.01	99.84	82.29
Ours	4.04 $\pm$ 3.01	84.08	99.86	82.91
▼ In-House Dataset				
RIoU3D (2020) [28]	57.36 $\pm$ 46.94	86.73	99.98	88.67
FCOS3D (2021) [21]	4.94 $\pm$ 3.84	84.22	99.78	81.44
GWD3D (2023) [23]	58.09 $\pm$ 45.29	86.62	99.99	87.53
PSC (2024) [25]	3.47 $\pm$ 2.89	87.03	99.94	91.02
Ours	3.05 $\pm$ 2.42	87.15	99.99	91.51
▼ V2S-Real Dataset				
RIoU3D (2020) [28]	62.21 $\pm$ 50.47	87.87	99.65	91.95
FCOS3D (2021) [21]	4.65 $\pm$ 3.57	89.07	99.62	94.71
GWD3D (2023) [23]	63.72 $\pm$ 47.10	88.06	99.75	90.59
PSC (2024) [25]	3.52 $\pm$ 2.61	89.14	99.70	96.42
Ours	2.75 $\pm$ 2.07	89.54	99.70	96.95

the same training settings (learning rate, batch size, optimizer, etc.) for fair comparison. Mean Intersection over Union (mIoU) and Average Precision (AP) are used for localization, and angle deviation for 3D orientation evaluation. All models use a ResNet-50 backbone [10], while retaining their original orientation representations and loss formulations. Training is performed with AdamW [14] for 36 epochs, with an initial learning rate of $1e-4$, 500-iteration warm-up, and $\times 0.1$ decay at epochs 24 and 33. Random rotation is used for data augmentation. Training and testing data are curated using our pipeline in Sec. 3.4.

Datasets. We combine all training sets to train a unified model for release, and evaluate it separately on the test sets of the four datasets: **1) ACDC [3].** It contains 150 SAX CMR CINE scans (100 training / 50 testing). End-diastole (ED) and end-systole (ES) frames are annotated with masks. **2) M&Ms-2 [6, 17].** This multi-center dataset includes 360 SAX CMR CINE scans with ED/ES annotations, with 160 scans held out for testing. **3) In-House.** It contains 188 CINE scans (95 for testing) annotated on all 30 frames, ensuring the trained model generalizes across arbitrary cardiac phases. **4) V2S-Real.** It contains 2,000 in-house isotropic 3D frames from 10 volunteers across 10 cardiac cycles, with the first frame manually annotated and the remaining frames propagated

Table 2. Ablation studies with the length of encoding and the weight of embedding loss on V2S-Real dataset.

N	Angle Dev.	AP_{75}	w_{emb}	Angle Dev.	AP_{75}
3	2.99 ± 2.15	96.13	0.5	3.27 ± 2.79	95.89
4	2.75 ± 2.07	96.95	1.0	2.91 ± 2.14	96.26
7	3.09 ± 2.60	94.68	2.0	2.75 ± 2.07	96.95
13	3.16 ± 2.67	93.63	5.0	3.15 ± 3.32	89.07

**Fig. 5.** Comparison on the real-world SAX plane prescription task. We compare our predicted SAX normal with manually prescribed planes, demonstrating performance on par with existing automatic multi-view plane prescription methods [15, 22, 8].

using [27]. The isotropic resolution avoids resampling degradation and enables realistic arbitrary slicing for evaluation.

Main results. Results are shown in Tab. 1, with visualization shown in Fig. 2. IoU- and Gaussian-based methods achieve competitive localization accuracy, but exhibit large angle deviations due to their inherent 180° rotational ambiguity. L1-based FCOS3D covers the full angular range, yet suffers from boundary discontinuity. PSC alleviates the in-plane boundary issue, but remains unstable as $z \rightarrow 1$. In contrast, our method resolves these limitations and achieves the best overall performance (averaged over four datasets): mIoU of 86.18% (0.35% higher than PSC), AP_{75} of 87.96% (0.89% higher), and angle deviation of 3.39° (0.65° lower). Owing to its continuous and stable representation over the full orientation range, it also yields lower standard deviation. Because both training and testing slices are generated through the same resampling pipeline, the model may partially exploit interpolation-related artifacts from anisotropic volumes rather than true anatomical cues. This risk is partially mitigated by the additional V2S-Real experiment, where isotropic acquisition avoids resampling degradation and demonstrates consistent generalization.

Ablation studies. The embedding length $N = 4$ in PCC and the weight of the embedding loss $w_{emb} = 2$ are determined according to Tab. 2.

Comparison with multi-view. We evaluate four volunteers in a real-world clinical prescription setting. Following the clinical workflow, localizers are first acquired, and pseudo 2CH/4CH views are subsequently captured for manual multi-view estimation of the SAX normal. The reference normal is derived from the SAX DICOM metadata of the clinical process, providing reliable ground

truth for comparison with our predictions (see Fig. 5a–b). The summarized angle deviations in Fig. 5c demonstrate the effectiveness of the proposed approach.

5 Conclusion

We present a novel paradigm for single-slice LV localization and 3D orientation estimation in CMR, offering a simpler alternative to view classification and geometric intersection approaches. To support this task, we develop a scalable data curation pipeline to build a dedicated benchmark. We adapt four representative orientation-aware detection frameworks and analyze their limitations when applied to SAX orientation estimation in CMR. To address these issues, we introduce a Polar-Coupled Circular (PCC) embedding mechanism that provides a continuous and unambiguous orientation representation. Extensive experiments on four datasets demonstrate strong performance, achieving an average mIoU of 86.18%, an angle deviation of 3.39° , and an AP_{75} of 87.96%.

Acknowledgments. This study was supported in part by The Ohio State University Clinical and Translational Science Institute (CTSI) and the National Center for Advancing Translational Sciences of the National Institutes of Health (NIH) under Grant UM1TR004548, and by NIH Grant R01 HL148103.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Aung, N., Lee, A.M., Sanghvi, M.M., Fung, K., Paiva, J.M., Thomson, R.J., et al.: Automatic classification of cmr image sequences with convolutional neural networks. *European Heart Journal - Cardiovascular Imaging* **20**(Supplement2), jez116.025 (2019)
2. Bağci, U., Udupa, J.K., Chen, X.: Orientation estimation of anatomical structures in medical images for object recognition. In: *Medical Imaging 2011: Image Processing*. vol. 7962, p. 79622L. International Society for Optics and Photonics (2011)
3. Bernard, O., Lalande, A., Zotti, C., Cervenansky, F., Yang, X., Heng, P.A., et al.: Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: Is the problem solved? *IEEE Transactions on Medical Imaging* **37**(11), 2514–2525 (2018)
4. Blansit, K., Retson, T., Masutani, E., Bahrami, N., Hsiao, A.: Deep learning–based prescription of cardiac mri planes. *Radiology: Artificial Intelligence* **1**(6), e180069 (2019)
5. Böttcher, B., Meinel, F.G., Deyerberg, K.K., Watzke, L.M., Manzke, M., Gorodezky, M., et al.: Fully automated plane prescription in cardiac mri: A prospective cohort study. *Journal of Magnetic Resonance Imaging* **63**(3), 891–903 (2026)
6. Campello, V.M., Gkontra, P., Izquierdo, C., Martín-Isla, C., Sojoudi, A., Full, P.M., et al.: Multi-centre, multi-vendor and multi-disease cardiac segmentation: The m&ms challenge. *IEEE Transactions on Medical Imaging* **40**(12), 3543–3554 (2021)

7. Chauhan, D., Anyanwu, E., Goes, J., Besser, S.A., Anand, S., Madduri, R., et al.: Comparison of machine learning and deep learning for view identification from cardiac magnetic resonance images. *Clinical Imaging* **82**, 121–126 (2022)
8. Edalati, M., Zheng, Y., Watkins, M.P., Chen, J., Liu, L., Zhang, S., et al.: Implementation and prospective clinical validation of ai-based planning and shimming techniques in cardiac mri. *Medical Physics* **49**(1), 129–143 (2022)
9. Glessgen, C., Crowe, L.A., Wetzl, J., Schmidt, M., Yoon, S.S., Vallée, J.P., Deux, J.F.: Automated vs manual cardiac mri planning: a single-center prospective evaluation of reliability and scan times. *European Radiology* **35**, 3927–3936 (2025)
10. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: *IEEE Conference on Computer Vision and Pattern Recognition*. pp. 770–778 (2016)
11. Kellman, P., Lu, X., Jolly, M.P., Bi, X., Kroeker, R., Schmidt, M., et al.: Automatic lv localization and view planning for cardiac mri acquisition. *Journal of Cardiovascular Magnetic Resonance* **13**(1), P39 (2011)
12. Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S.: Feature pyramid networks for object detection. In: *IEEE Conference on Computer Vision and Pattern Recognition*. pp. 2117–2125 (2017)
13. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **42**(2), 318–327 (2020)
14. Loshchilov, I., Hutter, F.: Decoupled weight decay regularization. In: *International Conference on Learning Representations* (2018)
15. Lu, X., Jolly, M.P., Georgescu, B., Hayes, C., Speier, P., Schmidt, M., et al.: Automatic view planning for cardiac mri acquisition. In: *International Conference on Medical Image Computing and Computer Assisted Intervention*. pp. 479–486 (2011)
16. Ma, X., Ouyang, W., Simonelli, A., Ricci, E.: 3d object detection from images for autonomous driving: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **46**(5), 3537–3556 (2024)
17. Martín-Isla, C., Campello, V.M., Izquierdo, C., Kushibar, K., Sendra-Balcells, C., Gkontra, P., et al.: Deep learning segmentation of the right ventricle in cardiac mri: The m&ms challenge. *IEEE Journal of Biomedical and Health Informatics* **27**(7), 3302–3313 (2023)
18. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., et al.: Pytorch: An imperative style, high-performance deep learning library. In: *Conference on Neural Information Processing Systems*. vol. 32, pp. 8024–8035 (2019)
19. Tian, Z., Shen, C., Chen, H., He, T.: Fcos: Fully convolutional one-stage object detection. In: *IEEE/CVF International Conference on Computer Vision*. pp. 9626–9635 (2019)
20. Vergani, V., Razavi, R., Puyol-Antón, E., Ruijsink, B.: Deep learning for classification and selection of cine cmr images to achieve fully automated quality-controlled cmr analysis from scanner to report. *Frontiers in Cardiovascular Medicine* **8** (2021)
21. Wang, T., Zhu, X., Pang, J., Lin, D.: Fcos3d: Fully convolutional one-stage monocular 3d object detection. In: *IEEE/CVF International Conference on Computer Vision Workshops*. pp. 913–922 (2021)
22. Wei, D., Ma, K., Zheng, Y.: Training automatic view planner for cardiac mr imaging via self-supervision by spatial relationship between views. In: *International Conference on Medical Image Computing and Computer Assisted Intervention*. pp. 526–536 (2021)

23. Yang, X., Zhang, G., Yang, X., Zhou, Y., Wang, W., Tang, J., et al.: Detecting rotated objects as gaussian distributions and its 3-d generalization. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **45**(4), 4335–4354 (2023)
24. Yu, Y., Da, F.: Phase-shifting coder: Predicting accurate orientation in oriented object detection. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2023)
25. Yu, Y., Da, F.: On boundary discontinuity in angle regression based arbitrary oriented object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **46**(10), 6494–6508 (2024)
26. Yu, Y., Martin, P., Bu, Z., Liu, Y., Zheng, Y.Y., Simonetti, O., Han, Y., Xue, Y.: Uncertainty-aware structured data extraction from full cmr reports via distilled llms. In: *IEEE 23rd International Symposium on Biomedical Imaging* (2026)
27. Zhang, Z., Yu, Y., Zhu, S., Aly, A., Gao, Y., Gu, N., Xue, Y.: Atlas is your perfect context: One-shot customization for generalizable foundational medical image segmentation. *arXiv Preprint arXiv:2512.18176* (2025)
28. Zheng, Y., Zhang, D., Xie, S., Lu, J., Zhou, J.: Rotation-robust intersection over union for 3d object detection. In: *European Conference on Computer Vision*. pp. 464–480 (2020)
29. Zhou, Y., Yang, X., Zhang, G., Wang, J., Liu, Y., Hou, L., et al.: Mmrotate: A rotated object detection benchmark using pytorch. In: *30th ACM International Conference on Multimedia*. pp. 7331–7334 (2022)