This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

An Anatomical Significance-Aware Architecture for Explainable Myocardial Infarction Prediction via Multi-Task Learning

Jiachuan Peng[1⋆[0000−0002−2209−0348]], Marcel Beetz[1], Abhirup Banerjee[1,2[0000−0001−8198−5128]], Min Chen[1[0000−0001−5320−5729]], and Vicente Grau[1[0000−0001−8139−3480]]

[1] Department of Engineering Science, University of Oxford, Oxford OX3 7DQ, UK
jiachuan.peng@seh.ox.ac.uk
[2] Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford OX3 9DU, UK

Abstract. Myocardial infarction (MI) is a significant health burden globally. Its precise prediction is critical yet complicated by the functional complexities of the heart and heterogeneous clinical presentations. Although learning-based methods that model the 3D heart anatomy have been widely studied, improving cardiac embeddings with localized substructures in a multi-task setting, remains under-explored. In this work, we present a novel deep learning model that produces explainable embeddings with high relevance to cardiac function via multi-task learning. Its transformer-based architecture contains modules for both MI classification and cardiac substructure prediction. By jointly learning these tasks with shared embeddings, the model is able to better capture 3D cardiac geometries and deformation across cardiac phases, enhancing its predictive ability. We evaluate the proposed method on cardiac anatomies captured during end-diastolic and end-systolic phases from the UK Biobank study. Compared to the existing learning-based benchmarks, our method exhibits high predictive performance, achieving an area under the receiver operating characteristic curve for MI prediction of 0.802. We also demonstrate the strong explainability of our model by showing that the latent features generated under the proposed multi-task setting have a strong and statistically significant correlation with key clinical markers, such as ejection fraction.

Keywords: Multi-task learning · Geometric deep learning · Myocardial infarction · Explainability · Point cloud transformer

1 Introduction

One leading group of causes of mortality in the world is cardiovascular diseases, among which myocardial infarction (MI) is one of the major threats [8]. Improving the prediction of MI events at an early age can enable targeted and

⋆ Corresponding author

preventive therapies, enhance the average survival rates, and ease the burden on healthcare systems. Therefore, there is a growing need to develop advanced predictive models that are capable of integrating cardiac pathophysiological details for precise MI prediction [21].

Numerous learning-based approaches have been used to extract key cardiac features, such as functional attributes [22, 24], heart shape [1] or contours of heart cavities [13, 14], for automated cardiac disease characterization. However, the predictive performance of these methods can be hindered by the limited information contained in handcrafted features or 2D segmentation maps. More recently, deep learning techniques have shown promise for directly analyzing high-resolution imaging data [12, 16, 23, 26, 27]. Of particular note, several geometry-based studies have built a prediction pipeline based on 3D heart data [3, 5, 7, 10]. Leveraging 3D anatomy reconstructed from 2D cine cardiac magnetic resonance (CMR) images, the full cardiac geometry and deformation can be modeled for disease prediction. In [4], the authors proposed a multi-task architecture to perform 3D anatomy modeling and MI prediction simultaneously. While the previous method seeks to explicitly model substructure information as the input for reconstruction, this design eases the difficulty of the substructure classification task, which may compromise cardiac feature learning.

In this paper, we propose a multi-task learning framework to enhance MI prediction via an interpretable cardiac substructure prediction task. We build a transformer-based network to learn clinically meaningful representations from multi-phase biventricular anatomy in 3D point clouds. The model processes the input data by first generating point patches and then building patch-to-patch dependencies via the attention mechanism. However, for MI prediction where data can be sparse, single-task training can yield sub-optimal and biased embeddings [9, 25]. To address this, an auxiliary substructure prediction branch is introduced to jointly supervise the training process. By identifying localized sub-regions of the heart, the model learns to mine anatomical details with greater diversity and variability in a fine-grained manner. Moreover, the shared embeddings under the multi-task setting demonstrate improved generalization and clinical reliability by integrating individual morphological details that complement the global MI prediction objective.

Our experimental results demonstrate that the proposed approach outperforms previous methods, setting a new state-of-the-art benchmark in MI prediction. Moreover, correlation analysis with key biomarkers related to cardiac functions showcases the model's ability to capture clinically relevant shape deformation across cardiac phases. Visualization of attention maps and latent space further supports that our multi-task design refines the transformer architecture's understanding of the clinical importance of different cardiac structures. As a result, the model learns to selectively attend to the most relevant regions for MI prediction, leading to more transparent and rational decision-making.

2 Materials and Methods

2.1 Dataset

We conduct our experiments on a dataset of 470 subjects collected from the UK Biobank study [19]. Half of the subjects developed MI after the imaging date (i.e., incident MI), while the other half were considered as healthy, with no prior history of diseases listed in [4]. For each subject, biventricular anatomies are captured at end-diastolic (ED) and end-systolic (ES) cardiac phases. These anatomies are represented as 3D point cloud surface data, reconstructed from cine CMR images following the protocols outlined in [2, 6]. We perform pre-processing by translating the point sets to a zero-mean coordinate system, and standardizing the scale using the standard deviation of the dataset. The dataset is split into 70%/5%/25% for training, validation, and testing.

2.2 Framework

As depicted in Fig. 1, we propose a dual-branch framework to extract clinically meaningful and discriminative features by leveraging localized information from cardiac substructures. It comprises modules for partitioning the input point cloud anatomies into patches and embedding them into higher-dimensional latent space, followed by a series of transformer blocks as the encoder to generate intermediate features. The substructure prediction module is devised for feature refinement by enhancing the awareness of spatial importance. Finally, the model training is guided by supervision at multiple levels of granularity. Details are given below.

Similar to [18, 28], we employ farthest point sampling (FPS) to split the point sets into patches. Concretely, given input anatomies $X \in \mathbb{R}^{N\times 3}$ with N three-dimensional points, FPS selects a subset of g centroid points that are maximally spaced to ensure uniform coverage of the geometry. The K-nearest neighbors (KNN) algorithm is then applied to compute the distances from the centers and select the k closest points based on minimum distance. This results in a structured representation of the anatomies as patches $P \in \mathbb{R}^{g\times k\times 3}$. We empirically find that the model reaches better convergence when $g \times k = 2N$.

The generated patches are projected into initial embeddings via a simplified PointNet [20] where each patch undergoes 1D convolution and max-pooling operations. Next, we apply positional encoding to embeddings. Formally, this results in the input sequence $\mathbf{z} = \{z_1, z_2, z_3, ..., z_g\}$ with a classification token $[cls]$ for encoding, which is initialized as a zero vector. As the information contained in individual patches is limited, we adopt vision transformer (ViT) [11] as the encoder to compute mutual dependencies across embeddings with multi-head self-attention. Mathematically, we compute the query Q, key K, and value V by performing a linear transformation to $\mathbf{z}$ with learnable weights. The attention is expressed as:

$$\text{attention} = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \tag{1}$$

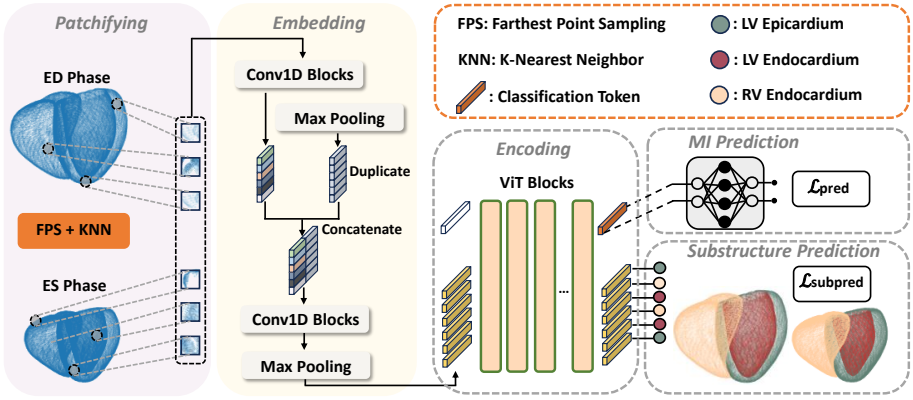

Fig. 1. Our proposed dual-branch framework for improved MI prediction. A patchifying module partitions the input ED and ES anatomies into point patches using FPS and KNN. Patches are then projected into initial embeddings, which are fed into the transformer encoder, together with an empty classification token. The substructure prediction module assigns each patch embedding to its corresponding cardiac substructure category, which encourages the model to focus on more localized spatial relationships, thereby refining the learned representations.

where d_k represents the dimension of K column-wise.

In this work we focus on three key cardiac substructures, due to their importance in MI prediction: left ventricle (LV) endocardium, LV epicardium, and right ventricle (RV) endocardium. In order to differentiate between these three substructures in the embedding space and thus enable more localized anatomical information to supervise the model training, we introduce a cardiac substructure prediction module. Normally, the model predicts the substructure category for each individual point. However, this can introduce redundancy, as point cloud representations are dense, and points within the same patch typically share the same category. Instead, we formulate the point-wise segmentation task as a patch-wise classification. Specifically, we attach a shared-weight multilayer perceptron (MLP) prediction head to each embedding in the sequence $\mathbf{z}$ and output the predicted substructure categories, which are compared with the ground-truths $\mathbf{s} = \{s_1, s_2, s_3, ..., s_g\}$. We define the substructure prediction loss $\mathcal{L}_{subpred}$ using cross-entropy as

$$\mathcal{L}_{subpred} = -\sum_{i=1}^{C} \mathbf{s}_i \log(MLP(\mathbf{z}))_i, \tag{2}$$

where $C = 3$ to represent the three substructures. The refined embeddings contain rich and fine-grained spatial information. The classification token $[cls]$ aggregates global information by attending to all other tokens in the sequence and accumulating contexts via self-attention. For MI prediction, we define the loss

$\mathcal{L}_{pred}$ which also adopts the cross-entropy loss, ensuring simplicity in weight balancing.

We define our multi-task training objective as a linear combination of two loss terms, given as:

$$\mathcal{L}_{total} = \mathcal{L}_{pred} + \mathcal{L}_{subpred} * \alpha, \tag{3}$$

where α is the coefficient to weight task importance.

2.3 Implementation Details and Metrics

In our experiments, the hyperparameters corresponding to the properties of the anatomies are $N = 36000$, $g = 480$, and $k = 150$. The patch embedding dimension is set to 384, and the weighting term α in Eq. (3) is set to 1. Our model employs 12 ViT blocks, with 6 attention heads and a dropout rate of 0.1. We optimize our model using the AdamW optimizer with an initial learning rate of 0.0001 for 70 epochs. During training, the learning rate is decreased based on the cosine annealing strategy. Our model is implemented by Pytorch and is trained on NVIDIA 2080Ti GPU with 11 GB of memory.

To quantify the MI prediction performance, we utilize the area under the receiver operating characteristic curve (AUROC), accuracy, and F1 score. They comprehensively assess the model's ability to distinguish between positive and negative classes under different threshold values.

3 Results and Discussion

3.1 Myocardial Infarction Prediction

To evaluate the effectiveness of the proposed multi-task learning model in MI prediction, we conduct a comparative study against several state-of-the-art learning-based approaches. These include: 1) standard CNN with 2D segmentation contours; 2) 3D Contraction [10], which leverages principle component analysis (PCA) to extract features from LV meshes for MI prediction; 3) Point Cloud Classification Network (PCCN) [7]; 4) Point Cloud Deformation Network (PCD-Net) [5], which explicitly models cardiac deformation within latent space; and 5) Multi-objective Point Cloud Autoencoder (MOPCA) [4]: a multi-task network for anatomy reconstruction and MI prediction.

Table 1 summarizes the quantitative results of MI prediction. It is observed that our proposed model substantially outperforms all other methods across all metrics. Notably, compared to MOPCA which also improves MI prediction via multi-task learning, our model achieves an approximately 5% performance gain, highlighting its superiority in learning cardiac representations.

3.2 Ablation Study

To assess the impact of multi-task learning in the proposed model, we perform an ablation study by varying the weighting coefficient α, i.e., the contribution of

Table 1. Comparison of MI prediction results. All results are reported as the mean values, with the highest highlighted in bold. The higher the value the better.

Methods	Input	Modality	AUROC	Accuracy	F1 score
CNN	LV + RV	2D contours	0.641	0.608	0.617
3D Contraction [10]	LV	mesh	0.647	0.614	0.605
PCCN [7]	LV + RV	point	0.646	0.652	0.637
PCD-Net [5]	LV + RV	point	0.721	0.668	0.663
MOPCA [4]	LV + RV	point	0.767	0.694	0.695
Ours	LV + RV	point	**0.802**	**0.735**	**0.730**

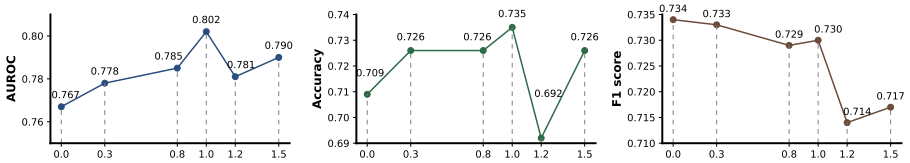

Fig. 2. MI prediction performance under different weighting coefficient α in multi-task learning.

the substructure prediction task, in the total loss function. The numerical results are illustrated in Fig. 2, reported according to the aforementioned metrics.

Generally, we find that prediction performance gradually improves as α increases, with the optimal trade-off reached at $\alpha = 1$. While we observe minor changes in the F1 score (0.734 at $\alpha = 0$ vs. 0.730 at $\alpha = 1$), both AUROC and accuracy show considerable growth, peaking at 0.802 and 0.735, respectively. However, further increasing α to values such as 1.2 and 1.5 results in performance drop. Given that both tasks utilize the same training objective and therefore have similar loss magnitudes, we hypothesize that the overemphasis on the substructure prediction task might bias the optimization process, hence degrading the MI prediction. Conversely, this finding indicates that our multi-task learning formulation is straightforward yet effective. It is also noteworthy that even without the substructure prediction task ($\alpha = 0$), the model can still perform slightly better than MOPCA. We credit this advantage to the design of the framework backbone. More importantly, the results support the efficacy of introducing auxiliary supervision in refining model training through more fine-grained annotations, which leads to enhanced MI prediction.

3.3 Clinical Relevance

In this section, we analyze the explainability of our proposed method by investigating the association between latent space and clinical patterns, both quantitatively and qualitatively. To quantify its relationship with cardiac function, we choose four widely used biomarkers: average myocardial thickness change

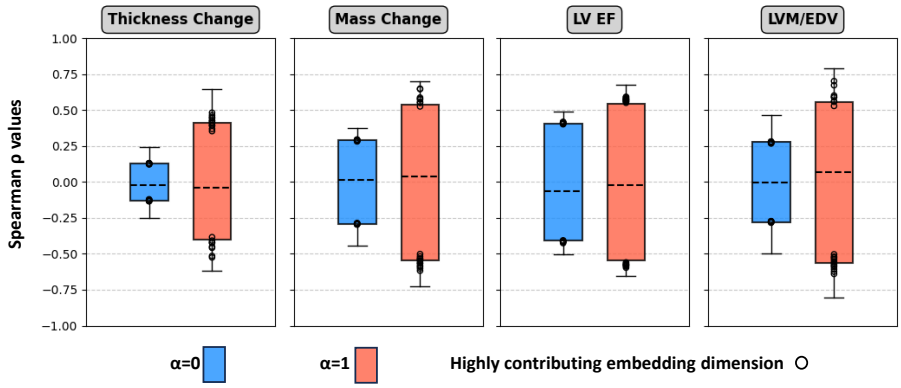

Fig. 3. Box plots of quantitative Spearman's rank correlation of each embedding dimension to 4 cardiac biomarkers: myocardial thickness change, mass change, LV ejection fraction (LVEF), and LV mass/end-diastolic volume ratio (LVM/EDV). Dots mark the correlation magnitude of the top 20 most influential embedding dimensions for predictions.

between ED and ES, myocardial mass change between ED and ES, LV ejection fraction (LVEF), and left ventricular mass to end-diastolic volume ratio (LVM/EDV). We compute Spearman's rank correlation coefficients ρ between each dimension of the generated embeddings (i.e., the classification token $[cls]$) and selected biomarkers, on the test data. All reported correlations are significant with $\mathbf{p} < 0.01$. Figure 3 summarizes the comparison of the results with ($\alpha = 1$) and without the substructure prediction task ($\alpha = 0$) in box plots. Additionally, we mark with dots the correlation magnitude of the top 20 embedding dimensions that contribute most to the model's predictions using SHAP [15]. In both cases, many of the top-ranked dimensions are found near the two extreme ends of the interquartile range (IQR), indicating that the model tends to utilize features with high clinical correlation primarily. A key distinction arises when the multi-task loss is used ($\alpha = 1$): in this case, the model produces embeddings with a broader correlation range, as evidenced by wider IQR and more dispersed distribution of top-ranked dimensions. In contrast, with the multi-task loss disabled ($\alpha = 0$), the dots become more concentrated and overlap closer to zero. This suggests that integrating the substructure prediction task allows the model to encode embeddings that are more clinically informative and diverse, while also improving its capacity to capture cardiac deformation (e.g., stronger LVEF correlation). Therefore, these refined embeddings enhance both model performance and interpretability of the decision-making process.

To further examine the effect of the substructure prediction task on the latent space of individual subjects, we visually compare the latent embeddings of each point patch alongside their corresponding attention maps. Specifically, the latent embeddings $\mathbf{z}$ are first obtained by passing point patches P through the transformer blocks. We then adopt UMAP [17] to reduce the dimension of each

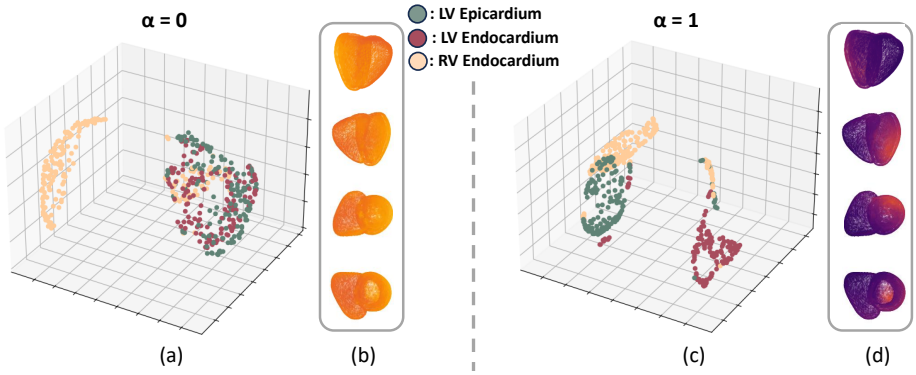

Fig. 4. UMAP visualization of latent embeddings $\mathbf{z}$ (a,c), with each point corresponding to one patch of the anatomy. Different colors indicate different substructure categories. The average attention maps are also visualized in three views alongside (b,d). The brighter the area, the greater the model's attention.

embedding to 3 for visualization in Fig. 4 (a,c). Additionally, the attention maps are generated by averaging the attention matrices over all transformer layers and attention heads (Fig. 4 (b,d)). Brighter areas indicate higher model attention. Figure 4 presents a representative example from a single subject for brevity; similar patterns are observed across all subjects in the dataset. Comparing the embedding space in (c) with (a), we observe that when $\alpha = 1$, embeddings of various substructures are well-clustered. Notably, the LV epicardium and endocaridum clusters are distinctly separated whereas in (a), they remain intertwined. Meanwhile, the LV endocardium is encapsulated within the epicardium, creating a thin interstitial space, i.e., myocardium, which is spatially subtle yet clinically critical for MI prediction. The enlarged degree of separation enables the model to amplify such indicative anatomical details, demonstrating the utility of our proposed model in capturing anatomical details from a finer level of granularity. Furthermore, we find that (b) focuses solely on the global view, failing to capture more detailed regions. In contrast, in (d), greater attention is directed to specific regions within the LV, particularly the anterior wall to the apex. This pattern appears to correspond to the clinical manifestation of MI, with the formation of localized injuries in the LV and subsequent remodeling. We believe the proposed module works as an optimization regularization term to direct attention to more clinically salient regions, eventually refining the learned cardiac representations.

4 Conclusion

In this paper, we present a transformer-based approach that leverages a multi-task learning strategy to jointly perform MI prediction and cardiac substructure classification. Our model encodes the 3D anatomy in the latent space, expressing a strong correlation to the functional and structural attributes of the heart, lead-

ing the model to concentrate on clinically informative phenotypes. The results highlight the effectiveness of introducing an auxiliary task with more detailed supervision to improve MI prediction with high interpretability. The proposed approach offers a promising direction for more robust and interpretable cardiac disease prediction.

Acknowledgments. This research has been conducted using the UK Biobank Resource under Application Number '40161'. The work of M. Beetz was supported by the Stiftung der Deutschen Wirtschaft (Foundation of German Business). The work of A. Banerjee was supported by the Royal Society University Research Fellowship under Grant URF\R1\221314. The work of V. Grau was supported by the CompBioMed 2 Centre of Excellence in Computational Biomedicine (European Commission Horizon 2020 research and innovation programme, grant agreement No. 823712).

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

References

1. Avard, E., Shiri, I., Hajianfar, G., Abdollahi, H., Kalantari, K.R., Houshmand, G., Kasani, K., Bitarafan-Rajabi, A., Deevband, M.R., Oveisi, M., et al.: Non-contrast cine cardiac magnetic resonance image radiomics features and machine learning algorithms for myocardial infarction detection. Computers in Biology and Medicine **141**, 105145 (2022)
2. Banerjee, A., Camps, J., et al.: A completely automated pipeline for 3D reconstruction of human heart from 2D cine magnetic resonance slices. Philosophical Transactions of the Royal Society A **379**(2212), 20200257 (2021)
3. Beetz, M., Acero, J.C., Banerjee, A., Eitel, I., Zacur, E., Lange, T., Stiermaier, T., Evertz, R., Backhaus, S.J., Thiele, H., et al.: Post-infarction risk prediction with mesh classification networks. In: International Workshop on Statistical Atlases and Computational Models of the Heart. pp. 291–301. Springer (2022)
4. Beetz, M., Banerjee, A., Grau, V.: Multi-objective point cloud autoencoders for explainable myocardial infarction prediction. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 532–542. Springer (2023)
5. Beetz, M., Banerjee, A., Grau, V.: Modeling 3D cardiac contraction and relaxation with point cloud deformation networks. IEEE Journal of Biomedical and Health Informatics (2024)
6. Beetz, M., Banerjee, A., Ossenberg-Engels, J., Grau, V.: Multi-class point cloud completion networks for 3D cardiac anatomy reconstruction from cine magnetic resonance images. Medical Image Analysis **90**, 102975 (2023)
7. Beetz, M., Yang, Y., Banerjee, A., Li, L., Grau, V.: 3D shape-based myocardial infarction prediction using point cloud classification networks. In: 2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). pp. 1–4. IEEE (2023)
8. Benjamin, E.J., Blaha, M.J., Chiuve, S.E., Cushman, M., Das, S.R., Deo, R., De Ferranti, S.D., Floyd, J., Fornage, M., Gillespie, C., et al.: Heart disease and stroke statistics—2017 update: a report from the american heart association. circulation **135**(10), e146–e603 (2017)

9. Caruana, R.: Multitask learning. Machine learning **28**, 41–75 (1997)
10. Corral Acero, J., Schuster, A., Zacur, E., Lange, T., Stiermaier, T., Backhaus, S.J., Thiele, H., Bueno-Orovio, A., Lamata, P., Eitel, I., et al.: Understanding and improving risk assessment after myocardial infarction using automated left ventricular shape analysis. Cardiovascular Imaging **15**(9), 1563–1574 (2022)
11. Dosovitskiy, A.: An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929 (2020)
12. Gupta, V., Petursson, P., Rawshani, A., Boren, J., Ramunddal, T., Bhatt, D.L., Omerovic, E., Angerås, O., Smith, G., Sattar, N., et al.: End-to-end deep-learning model for the detection of coronary artery stenosis on coronary ct images. Open Heart **12**(1) (2025)
13. Isensee, F., Jaeger, P.F., Full, P.M., Wolf, I., Engelhardt, S., Maier-Hein, K.H.: Automatic cardiac disease assessment on cine-mri via time-series segmentation and domain specific features. In: Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges: 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Quebec City, Canada, September 10-14, 2017, Revised Selected Papers 8. pp. 120–129. Springer (2018)
14. Khened, M., Alex, V., Krishnamurthi, G.: Densely connected fully convolutional network for short-axis cardiac cine mr image segmentation and heart diagnosis using random forest. In: Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges: 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Quebec City, Canada, September 10-14, 2017, Revised Selected Papers 8. pp. 140–151. Springer (2018)
15. Lundberg, S.: A unified approach to interpreting model predictions. arXiv preprint arXiv:1705.07874 (2017)
16. Luo, C., Shi, C., Li, X., Wang, X., Chen, Y., Gao, D., Yin, Y., Song, Q., Wu, X., Zhou, J.: Multi-task learning using attention-based convolutional encoder-decoder for dilated cardiomyopathy cmr segmentation and classification. Computers, Materials & Continua **63**(2) (2020)
17. McInnes, L., Healy, J., Melville, J.: Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426 (2018)
18. Pang, Y., Wang, W., Tay, F.E., Liu, W., Tian, Y., Yuan, L.: Masked autoencoders for point cloud self-supervised learning. In: European conference on computer vision. pp. 604–621. Springer (2022)
19. Petersen, S.E., Matthews, P.M., Francis, J.M., Robson, M.D., Zemrak, F., Boubertakh, R., Young, A.A., Hudson, S., Weale, P., Garratt, S., et al.: Uk biobank's cardiovascular magnetic resonance protocol. Journal of cardiovascular magnetic resonance **18**(1), 8 (2016)
20. Qi, C.R., Su, H., Mo, K., Guibas, L.J.: Pointnet: Deep learning on point sets for 3d classification and segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 652–660 (2017)
21. Reza-Soltani, S., Alam, L.F., Debellotte, O., Monga, T.S., Coyalkar, V.R., Tarnate, V.C.A., Ozoalor, C.U., Allam, S.R., Afzal, M., Shah, G.K., et al.: The role of artificial intelligence and machine learning in cardiovascular imaging and diagnosis. Cureus **16**(9) (2024)
22. Sanchez-Martinez, S., Duchateau, N., Erdei, T., Kunszt, G., Aakhus, S., Degiovanni, A., Marino, P., Carluccio, E., Piella, G., Fraser, A.G., et al.: Machine learning analysis of left ventricular function to characterize heart failure with preserved ejection fraction. Circulation: cardiovascular imaging **11**(4), e007138 (2018)

23. Snaauw, G., Gong, D., Maicas, G., Van Den Hengel, A., Niessen, W.J., Verjans, J., Carneiro, G.: End-to-end diagnosis and segmentation learning from cardiac magnetic resonance imaging. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). pp. 802–805. IEEE (2019)
24. Tabassian, M., Sunderji, I., Erdei, T., Sanchez-Martinez, S., Degiovanni, A., Marino, P., Fraser, A.G., D'hooge, J.: Diagnosis of heart failure with preserved ejection fraction: machine learning of spatiotemporal variations in left ventricular deformation. Journal of the American society of echocardiography **31**(12), 1272–1284 (2018)
25. Thrun, S.: Is learning the n-th thing any easier than learning the first? Advances in neural information processing systems **8** (1995)
26. Xia, C., Li, X., Wang, X., Kong, B., Chen, Y., Yin, Y., Cao, K., Song, Q., Lyu, S., Wu, X.: A multi-modality network for cardiomyopathy death risk prediction with cmr images and clinical information. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 577–585. Springer (2019)
27. Yang, G., Chen, J., Sheng, X., Yang, S., Zhuang, X., Raman, B., Li, L., Grau, V.: Contrast-free myocardial scar segmentation in cine mri using motion and texture fusion. arXiv preprint arXiv:2501.05241 (2025)
28. Yu, X., Tang, L., Rao, Y., Huang, T., Zhou, J., Lu, J.: Point-bert: Pre-training 3d point cloud transformers with masked point modeling. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 19313–19322 (2022)