Coarse-Grained Mask Regularization for Microvascular Obstruction Identification from non-contrast Cardiac Magnetic Resonance

Yige Yan[1⋆], Jun Cheng[2(✉)], Xulei Yang[2], Zaiwang Gu[2], Shuang Leng[3,4], Ru San Tan[3,4], Liang Zhong[3,4], and Jagath C. Rajapakse[1(✉)]

[1] College of Computing and Data Science, Nanyang Technological University, Singapore
asjagath@ntu.edu.sg
[2] Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), 1 Fusionopolis Way, Connexis, 138632, Singapore
cheng_jun@i2r.a-star.edu.sg
[3] National Heart Centre Singapore, 5 Hospital Drive, 169609, Singapore
[4] Duke-NUS Medical School, National University of Singapore, Singapore
gmszl@nus.edu.sg

Abstract. Identification of microvascular obstruction (MVO) in acute myocardial infarction patients is critical for prognosis and has a direct link to mortality risk. Current approaches using late gadolinium enhancement (LGE) for contrast-enhanced cardiovascular magnetic resonance (CMR) pose risks to the kidney and may not be applicable to many patients. This highlights the need to explore alternative non-contrast imaging methods, such as cine CMR, for MVO identification. However, the scarcity of datasets and the challenges in annotation make the MVO identification in cine CMR challenging and remain largely under-explored. For this purpose, we propose a non-contrast MVO identification framework in cine CMR with a novel coarse-grained mask regularization strategy to effectively utilize information from LGE annotations in training. We train and validate our model on a dataset comprising 680 cases. Our model demonstrates superior performance over competing methods in cine CMR-based MVO identification, proving its feasibility and presenting a novel and patient-friendly approach to the field. The code is available at https://github.com/code-koukai/MVO-identification.

Keywords: Microvascular Obstruction · Non-Contrast MRI · Spatio-temporal Features.

1 Introduction

Acute myocardial infarction (AMI) is a leading cause of mortality globally [3,4], and among its complications, microvascular obstruction (MVO) is particularly concerning. MVO refers to the blockage of small blood vessels within the heart,

⋆ The work was done during an internship at I^{2}R, A*STAR.

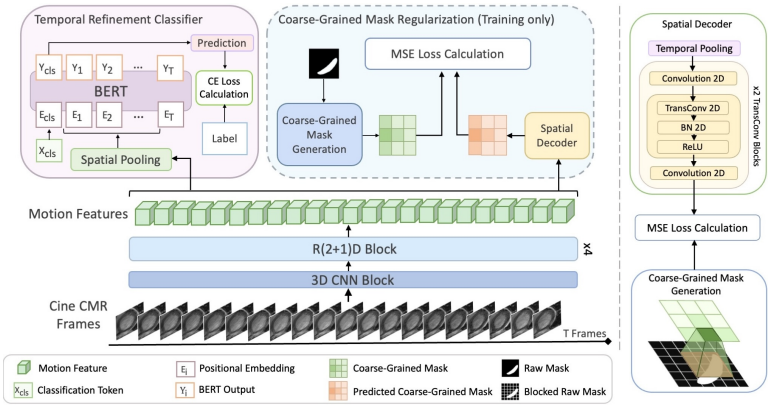
(a) Cine CMR (b) LGE CMR

Fig. 1: The illustration of cine CMR and LGE CMR: (a) Cine CMR sequences and a frame in diastolic phase; (b) The LGE CMR image corresponding to the frame in diastolic phase in (a) and its MVO region.

often occurring even after the main coronary arteries are reopened. This condition indicates significant cardiac damage and is associated with higher mortality and complications in AMI patients [8,1]. Clinically, MVO is significant due to its prediction of adverse outcomes, including extensive myocardial injury, poorer recovery, and higher rates of heart failure. Therefore, accurate identification of MVO is essential for effective case management and patient care. The current gold standard for detecting MVO is late gadolinium enhancement (LGE) cardiac magnetic resonance (CMR), which uses gadolinium to enhance contrast in CMR images, as shown in Fig. 1b, providing critical information for guiding clinical management of AMI patients [14,22].

Several studies have explored deep-learning techniques for MVO identification from LGE images [15,6]. However, the use of LGE poses a risk to kidneys, and studies indicate that around 20% of AMI patients are unable to undergo LGE due to severe renal impairment [23]. Moreover, incorporating LGE into CMR imaging leads to increased scanning time and treatment expenses [30]. Therefore, there is a pressing need to minimize gadolinium-based contrast agents in cardiac imaging for MVO identification.

Efforts have been made to leverage cine CMR as a contrast-free method for detecting myocardial lesions by observing regional wall motion abnormalities [16]. Cine CMR, as shown in Fig. 1a, is commonly used for cardiac reconstruction tasks [20,18]. It can also serve as an auxiliary modality for myocardial tissue segmentation [28,17]. Recent studies have begun to explore the detection of cardiac tissue and lesions in cine CMR images [29,10]. The importance of spatiotemporal features for cardiac lesion detection is highlighted by various approaches [12,19,26,27,31]. Although cine CMR has made strides in detecting some specific myocardial lesions, its potential for MVO remains largely under-explored. This difficulty mainly stems from data scarcity and challenges in annotating MVO in cine CMR sequences. Alternatively, we are able to delineate MVO regions from LGE imaging and align them to cine CMR sequences. However, this approach is prone to misalignment errors, and such errors significantly impair the delineation accuracy. This is especially critical since MVO regions typically have small sizes, where even slight errors can lead to an exaggerated impact. Nevertheless, the presence of MVO can be determined accurately from LGE, which can serve as ground truth for model training.



(a) (b)

Fig. 2: Illustration of our proposed method includes (a) overview of our framework and (b) details of coarse-grained mask regularization. Initially, cine CMR images are processed with a 3D CNN block. Then we extract motion features using four R(2+1)D blocks and apply both spatial and temporal pooling. The spatially and temporally pooled features are input into BERT and a coarse-grained mask regularization branch to improve MVO feature learning. The dashed block is removed during the inference stage. The final prediction is obtained from BERT's output Y_{cls}.

To this end, we utilize LGE data to determine the presence of MVO and introduce a novel framework to achieve non-contrast MVO identification in AMI patients. Our approach incorporates a novel coarse-grained mask regularization strategy in training. The coarse-grained mask is derived from raw masks labelled in the LGE data, which might not perfectly align with MVO in cine CMR but are closely related to the actual MVO region. We then use this coarse-grained mask to regularize the model training. Inspired by spatiotemporal feature processing abilities of residual 2D+1D network (R(2+1)D) [25] and bidirectional encoder representations from transformers (BERT) [9,13], our network adopts R(2+1)D network for motion feature extraction and BERT for temporal refinement classifier in identifying MVO from non-contrast cine CMR. The proposed coarse-grained mask regularization is trained alongside the temporal refinement classifier using multi-task learning, as depicted in Fig. 2. It includes a spatial decoder that encourages the model to focus more on the probable MVO region, thus generating a common representation and alleviating overfitting in the identification task. Importantly, the regularization branch is removed after training, requiring only non-contrast cine CMR for inference. Our study suggests that the coarse-grained mask regularization enhances model training and improves MVO identification from non-contrast cine CMR. Our contributions can be summarized as follows:

1. We propose a novel approach for MVO identification by extracting spatiotemporal features from non-contrast cine CMR, which has not been well-explored previously;
2. To address the challenges in model training, we introduce a coarse-grained mask regularization strategy to leverage information from LGE data;
3. Our experimental results demonstrate the feasibility of identifying MVO from non-contrast cine CMR, potentially bridging the gap left by existing approaches that rely on harmful contrast agents.

2 Method

We propose a novel framework for MVO identification using non-contrast cine CMR. Our framework employs R(2+1)D motion feature extraction (Section 2.1) to compute spatiotemporal features. Further enhancement is achieved through a multi-task learning approach, which simultaneously integrates a temporal refinement classifier (Section 2.2) and coarse-grained mask regularization (Section 2.3), as illustrated in Fig. 2. An overall loss function is computed to effectively train the network (Section 2.4).

2.1 Motion Feature Extraction

Since MVO often manifests as localized motion abnormalities in small regions, effective motion feature extraction is essential for optimizing model performance. Drawing inspiration from R(2+1)D's ability to decompose 3D convolutions into (2+1)D filters, we employ this technique for motion feature extraction in MVO. This approach reduces model parameters, which is essential given the limited volume of MVO data. Additionally, it enables the accurate capture of spatial-temporal details in cine CMR sequences. Therefore, the utilization of R(2+1)D provides a balanced method that alleviates overfitting and ensures the detailed extraction of features from MVO data.

Our implementation starts with a 3D Convolutional Neural Network (CNN) block, providing efficient initialization to extract key spatiotemporal information. We then forward the processed feature to four R(2+1)D blocks for advanced motion feature extraction.

2.2 Temporal Refinement Classifier for MVO Identification

For the complex and subtle motion features of MVO, it is essential to use advanced temporal modeling for refinement. As illustrated in Fig. 2a, we adopt BERT for two main purposes: first, to capture temporal dependencies for feature refinement, and second, as a classifier for MVO identification.

After processing motion features from R(2+1)D blocks through spatial pooling, we fit them into BERT. The R(2+1)D blocks initially extract spatiotemporal features by separating 3D convolutions into spatial and temporal components,

thereby improving efficiency and performance. We further incorporate a classification embedding x_{cls} to learn the global context of the sequence. The input is then combined with learned positional embeddings E_i, which embed temporal order by representing the i^{th} position, enhancing the model's ability to understand sequence timing. Specifically, the positional embeddings E_i are critical for maintaining the temporal structure within BERT, ensuring accurate temporal context comprehension.

Furthermore, the BERT model fine-tunes these combined features to capture nuanced temporal dependencies critical for identifying MVO. Finally, the output Y_{cls} is derived, which serves as the classification result indicating the presence or absence of MVO. This structured approach, combining R(2+1)D for initial feature extraction and BERT for temporal refinement and classification, ensures robust handling of the subtle motion features of MVO.

2.3 Coarse-Grained Mask Regularization

With extracted features from R(2+1)D and BERT, it is intuitive to train a network to predict the MVO region. As we mentioned in the introduction, the challenge of this task is significantly amplified by the difficulty in obtaining accurate pixel-wise annotations. However, the presence of MVO can be determined accurately from LGE. As an initial study, we are working with a relatively small dataset. The current challenge is to find an effective way to train the model with this limited amount of data. To address this, we propose the coarse-grained mask regularization, as illustrated in Fig. 2b. This approach aims to empower the model to identify MVO presence by fully leveraging LGE annotations, integrating MVO region estimation as an auxiliary task to enhance utilization. Direct use of raw masks leads to decreased model performance. In contrast, our coarse-grained mask regularization converts pixel-wise yet misaligned location information into a block-wise regional context, thus minimizing the impact of misalignment. Simultaneously, this approach removes the reliance on defining quantitative thresholds to determine the presence of MVO within a single block. It shall be noted that the regularization branch is eliminated after the training process. As a result, we only need non-contrast cine CMR at the inference stage.

Coarse-Grained Mask Computation For a raw mask G with $n \times n$ pixels, we partition it into an $m \times m$ grid of non-overlapping blocks. Each block is denoted by G_{ij} with i and j representing the row and column indices respectively. We then introduce a computation for the MVO pixel ratio p_{ij}, defined as:

$$p_{ij} = \frac{\sum_{k=1}^{\lceil n/m \rceil} \sum_{l=1}^{\lceil n/m \rceil} G_{ij}(k,l)}{\lceil n/m \rceil^2}, \tag{1}$$

where $G_{ij}(k,l) = 1$ if the pixel at position (k,l) within the block G_{ij} is from an MVO region, and 0 otherwise. For edge cases where n is not evenly divisible by m, zero padding is applied to the remaining insufficient pixels to form a complete block. Finally, we derive an $m \times m$ matrix $P = \{p_{ij}\}$, providing a lower-resolution but insightful representation of the MVO regions.

Table 1: Performance comparison for competing methods

Method	AUC	Specificity	Recall	F1-score	Nb. parameters
ViViT[2]	58.54	67.44	45.45	43.48	41.98M
TimeSformer[5]	59.20	65.12	52.27	47.42	76.54M
C3D[24]	63.31	59.30	56.82	48.07	63.32M
I3D[7]	57.80	62.79	54.54	48.00	12.55M
SlowFast[11]	64.32	70.93	54.54	51.61	33.56M
P3D[21]	63.74	73.26	50.00	49.44	65.68M
Proposed Method	**72.70**	**75.58**	**56.82**	**55.56**	43.72M

Coarse-Grained MVO Prediction To utilize the coarse-grained mask obtained previously, we designed a spatial decoder to predict the MVO region, processing the features outlined in Section 2.1. As depicted in Fig. 2b, the decoder initiates with a temporal pooling layer, followed by two transposed convolution blocks. Each block comprises a 1×1 convolutional layer, a batch normalization layer, a maximum pooling layer, and a 3×3 transposed convolution layer. This configuration transforms the 3D feature maps into high-quality 2D predictions. Subsequently, we employ a mean-square-error (MSE) loss function to quantify the discrepancy between the predicted and actual coarse-grained MVO regions.

2.4 Overall Loss Function

The coarse-grained mask regularization branch is concurrently trained with the temporal refinement classifier branch through multi-task learning. To facilitate the simultaneous optimization of these tasks, we employ a loss function that combines cross-entropy (CE) loss L_{ce} for the classifier branch and MSE loss L_{mse} for the regularization branch. The formula is denoted as:

$$L_{overall} = L_{\text{ce}}(\hat{Y}, Y) + \lambda \cdot L_{\text{mse}}(\hat{P}, P), \tag{2}$$

where $\hat{Y}$ denotes the predicted logits, providing an assessment of MVO presence, and Y indicates the ground truth of MVO presence. $\hat{P}$ represents the predicted coarse-grained mask, as mentioned in Section 2.3, and P stands for the coarse-grained mask. λ controls the balance between the two branches.

3 Experimental Results

3.1 Dataset and Evaluation Metric

We collected short-axis cine and LGE CMR images on a 3T Siemens CMR scanner, resulting in a dataset of 680 cases (625 males; 57±9 years) with all images covering the entire left ventricle. Each case comprised a 30-frame cine sequence and a corresponding LGE image obtained during diastolic phases. The cine sequence included the entire cardiac cycle. Image analysis was performed in a core laboratory setting, where an expert with >9 years of experience in CMR

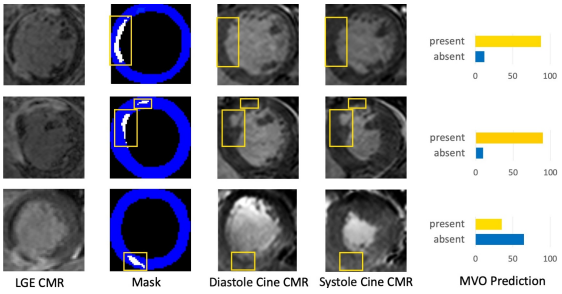


Fig. 3: Samples for visualization. The first and the second samples demonstrate successful MVO identification cases, while the last one shows a failed case. In LGE, the MVOs appear as dark spots in light myocardial areas. In Mask, blue denotes the myocardium and white shows MVO annotated based on LGE. Yellow boxes highlight MVO regions for clarity. MVO Prediction shows our model's confidence in each case.

Table 2: Ablation study of BERT and coarse-grained mask regularization

Components		Metrics				
BERT	CGMR	AUC	Specificity	Recall	F1-score	Nb. parameters
		64.45	69.77	45.45	44.44	31.50M
✓		65.45	70.93	56.82	53.91	37.86M
	✓	67.97	75.58	54.54	53.93	37.56M
✓	✓	**72.70**	**75.58**	**56.82**	**55.56**	43.72M

manually delineated both the endo- and epicardium after rigid registration of cine and LGE images using the freely available software Segment (version 4.0 R12067). The MVOs were extracted from the myocardium manually or using a semi-automatic approach on LGE images, serving as raw masks. The data were split into training (n=550, 221 MVO+) and testing (n=130, 44 MVO+). To reduce background noise and emphasize MVO movement, we cropped cine images and the masks to 64×64, ensuring the entire myocardium is presented in each image. We assessed our framework's performance using four standard metrics: Area Under Curve (AUC), Specificity, Recall, and F1-score.

3.2 Implementation Details

Cine CMR sequences underwent normalization and data augmentation in the training phase, including random cropping, flipping, rotation, and photometric distortions. The batch size was set to 16. We used the AdamW optimizer for models involving BERT, with a learning rate of 0.0001, and SGD for other models, with a learning rate of 0.01. A polynomial learning rate scheduler was employed to adjust the learning rate over 8000 iterations. Models were trained on an 80G NVIDIA A100 Tensor Core GPU on PyTorch version 1.7.1.

Table 3: Ablation study for the resolution of coarse-grained mask

Resolution	AUC	Specificity	Recall	F1-score
4×4	68.68	69.77	56.82	52.63
16×16	**72.70**	**75.58**	**56.82**	**55.56**
32×32	67.02	68.60	54.54	50.52
raw mask	62.23	69.77	45.45	44.44

Table 4: Ablation study for different λ in the loss function

value of λ	AUC	Specificity	Recall	F1-score
0.01	67.79	73.26	52.27	51.11
0.05	**72.70**	**75.58**	**56.82**	**55.56**
0.1	70.30	69.77	56.82	52.62

3.3 Performance Comparison

Given the largely unexplored nature of MVO identification in non-contrast imaging, we conduct performance comparisons with algorithms that undertake similar tasks to access our approach. We compare our approach with several competitive algorithms, including ViViT [2], TimeSformer [5], C3D [24], I3D [7], SlowFast [11], and P3D [21]. Experimental results presented in Table 1 demonstrate that our proposed model yields the best performance with an AUC of 72.70 and a specificity of 75.58, significantly outperforming other competing methods. Figure 3 reveals our model's precise identification of MVOs in low-motion situations. However, minor myocardial movements may obscure MVOs, making them challenging to differentiate from cardiac movements without MVO.

3.4 Ablation Study

Effectiveness of BERT and Coarse-Grained Mask Regularization We performed ablation studies on each component to evaluate their impact on overall performance. Our experiments included: the backbone alone, the backbone with BERT, the backbone with coarse-grained mask regularization, and the full framework. The results in Table 2 indicate that BERT for temporal refinement classifier and coarse-grained mask regularization, abbreviated as CGMR, both improve the model's performance.

Performance at Different Course-Grained Mask Resolutions We investigated how a coarse-grained mask impacts model performance across various resolutions, conducting ablation studies at 4×4, 16×16, 32×32 and the raw mask. Table 3 demonstrates resolution's significant effect on the model's performance. Additionally, using the raw mask lowers the performance, highlighting our method's effectiveness with annotation inaccuracies.

Performance for Different λ in the Loss Function We examined the impact of the loss weight ratio λ as referenced in 2.4 during model training. The λ was set between 0 and 1 to investigate performance discrepancies. Results in Table 4 illustrate that selecting an optimal λ can enhance the model's performance.

4 Conclusion

This study introduces a novel framework for MVO identification in non-contrast cine CMR by utilizing motion features and a unique coarse-grained mask regularization strategy. This method validates the feasibility of identifying MVO without the need for contrast agents and offers a more patient-friendly diagnostic process. Our work brings new perspectives for MVO identification using non-contrast CMR imaging. Limited by our data, our research focuses on the presence of MVO from cine CMR. Future work could explore identifying MVO regions from cine CMR.

Acknowledgments. This work is supported by the Agency for Science, Technology and Research under its AI^3 Horizontal Technology Coordinating Office Grant C231118001.

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

References

1. Amyar, A., Nakamori, S., Morales, M., Yoon, S., Rodriguez, J., Kim, J., Judd, R.M., Weinsaft, J.W., Nezafat, R.: Gadolinium-free cardiac mri myocardial scar detection by 4d convolution factorization. In: MICCAI. pp. 639–648 (2023)
2. Arnab, A., Dehghani, M., Heigold, G., Sun, C., Lučić, M., Schmid, C.: Vivit: A video vision transformer. In: ICCV. pp. 6836–6846 (2021)
3. Barnes, M., Heywood, A.E., Mahimbo, A., Rahman, B., Newall, A.T., Macintyre, C.R.: Acute myocardial infarction and influenza: a meta-analysis of case–control studies. Heart **101**(21), 1738–1747 (2015)
4. Beetz, M., Banerjee, A., Grau, V.: Multi-objective point cloud autoencoders for explainable myocardial infarction prediction. In: ICCV. pp. 532–542 (2023)
5. Bertasius, G., Wang, H., Torresani, L.: Is space-time attention all you need for video understanding? In: ICML. vol. 2, p. 4 (2021)
6. Brahim, K., Qayyum, A., Lalande, A., Boucher, A., Sakly, A., Meriaudeau, F.: A 3d deep learning approach based on shape prior for automatic segmentation of myocardial diseases. In: IPTA. pp. 1–6 (2020)
7. Carreira, J., Zisserman, A.: Quo vadis, action recognition? a new model and the kinetics dataset. In: CVPR. pp. 6299–6308 (2017)
8. De Waha, S., Patel, M.R., Granger, C.B., Ohman, E.M., Maehara, A., Eitel, I., Ben-Yehuda, O., Jenkins, P., Thiele, H., Stone, G.W.: Relationship between microvascular obstruction and adverse events following primary percutaneous coronary intervention for st-segment elevation myocardial infarction: an individual patient data pooled analysis from seven randomized trials. European heart journal **38**(47), 3502–3510 (2017)
9. Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding. NAACL pp. 4171–4186 (2019)
10. Dong, S., Pan, Z., Fu, Y., Yang, Q., Gao, Y., Yu, T., Shi, Y., Zhuo, C.: Deu-net 2.0: Enhanced deformable u-net for 3d cardiac cine mri segmentation. Medical Image Analysis **78**, 102389 (2022)

11. Feichtenhofer, C., Fan, H., Malik, J., He, K.: Slowfast networks for video recognition. In: ICCV. pp. 6202–6211 (2019)
12. Gonzales, R.A., Lamy, J., Seemann, F., Heiberg, E., Onofrey, J.A., Peters, D.C.: Tvnet: Automated time-resolved tracking of the tricuspid valve plane in mri long-axis cine images with a dual-stage deep learning pipeline. In: MICCAI. pp. 567–576 (2021)
13. Kalfaoglu, M.E., Kalkan, S., Alatan, A.A.: Late temporal modeling in 3d cnn architectures with bert for action recognition. In: ECCV 2020 Workshops. pp. 731–747 (2020)
14. Kim, R.J., Wu, E., Rafael, A., Chen, E.L., Parker, M.A., Simonetti, O., Klocke, F.J., Bonow, R.O., Judd, R.M.: The use of contrast-enhanced magnetic resonance imaging to identify reversible myocardial dysfunction. New England Journal of Medicine **343**(20), 1445–1453 (2000)
15. de La Rosa, E., Sidibé, D., Decourselle, T., Leclercq, T., Cochet, A., Lalande, A.: Myocardial infarction quantification from late gadolinium enhancement mri using top-hat transforms and neural networks. Algorithms **14**(8), 249 (2021)
16. Leiner, T.: Deep learning for detection of myocardial scar tissue: Goodbye to gadolinium? Radiology **291**(3), 618–619 (2019)
17. Li, L., Wu, F., Wang, S., Luo, X., Martín-Isla, C., Zhai, S., Zhang, J., Liu, Y., Zhang, Z., Ankenbrand, M.J., et al.: Myops: A benchmark of myocardial pathology segmentation combining three-sequence cardiac magnetic resonance images. Medical Image Analysis **87**, 102808 (2023)
18. Lyu, J., Li, G., Wang, C., Qin, C., Wang, S., Dou, Q., Qin, J.: Region-focused multi-view transformer-based generative adversarial network for cardiac cine mri reconstruction. Medical Image Analysis **85**, 102760 (2023)
19. Meng, Q., Bai, W., Liu, T., O'Regan, D.P., Rueckert, D.: Mesh-based 3d motion tracking in cardiac mri using deep learning. In: MICCAI. pp. 248–258 (2022)
20. Oksuz, I., Clough, J.R., Ruijsink, B., Anton, E.P., Bustin, A., Cruz, G., Prieto, C., King, A.P., Schnabel, J.A.: Deep learning-based detection and correction of cardiac mr motion artefacts during reconstruction for high-quality segmentation. TMI **39**(12), 4001–4010 (2020)
21. Qiu, Z., Yao, T., Mei, T.: Learning spatio-temporal representation with pseudo-3d residual networks. In: ICCV. pp. 5533–5541 (2017)
22. Reimer, K.A., Lowe, J.E., Rasmussen, M.M., Jennings, R.B.: The wavefront phenomenon of ischemic cell death. 1. myocardial infarct size vs duration of coronary occlusion in dogs. Circulation **56**(5), 786–794 (1977)
23. Shroff, G.R., Frederick, P.D., Herzog, C.A.: Renal failure and acute myocardial infarction: clinical characteristics in patients with advanced chronic kidney disease, on dialysis, and without chronic kidney disease. American heart journal **163**(3), 399–406 (2012)
24. Tran, D., Bourdev, L., Fergus, R., Torresani, L., Paluri, M.: Learning spatiotemporal features with 3d convolutional networks. In: ICCV. pp. 4489–4497 (2015)
25. Tran, D., Wang, H., Torresani, L., Ray, J., LeCun, Y., Paluri, M.: A closer look at spatiotemporal convolutions for action recognition. In: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. pp. 6450–6459 (2018)
26. Tripathi, P.C., Suvon, M.N., Schobs, L., Zhou, S., Alabed, S., Swift, A.J., Lu, H.: Tensor-based multimodal learning for prediction of pulmonary arterial wedge pressure from cardiac mri. MICCAI (2023)
27. Vimalesvaran, K., Uslu, F., Zaman, S., Galazis, C., Howard, J., Cole, G., Bharath, A.A.: Detecting aortic valve pathology from the 3-chamber cine cardiac mri view. In: MICCAI. pp. 571–580 (2022)

28. Wang, K.N., Yang, X., Miao, J., Li, L., Yao, J., Zhou, P., Xue, W., Zhou, G.Q., Zhuang, X., Ni, D.: Awsnet: an auto-weighted supervision attention network for myocardial scar and edema segmentation in multi-sequence cardiac magnetic resonance images. Medical Image Analysis **77**, 102362 (2022)
29. Xu, C., Xu, L., Ohorodnyk, P., Roth, M., Chen, B., Li, S.: Contrast agent-free synthesis and segmentation of ischemic heart disease images using progressive sequential causal gans. Medical image analysis **62**, 101668 (2020)
30. Yan, C., Hu, J., Li, Y., Xie, X., Zou, Z., Deng, Q., Zhou, X., Bi, X., Zeng, M., Liu, J.: Motion-corrected free-breathing late gadolinium enhancement combined with a gadolinium contrast agent with a high relaxation rate: an optimized cardiovascular magnetic resonance examination protocol. Journal of International Medical Research **48**(10), 0300060520964664 (2020)
31. Zhang, N., Yang, G., Gao, Z., Xu, C., Zhang, Y., Shi, R., Keegan, J., Xu, L., Zhang, H., Fan, Z., et al.: Deep learning for diagnosis of chronic myocardial infarction on nonenhanced cardiac cine mri. Radiology **291**(3), 606–617 (2019)