


BRIEF REPORT

Open Access



Automated cardiovascular MR myocardial scar quantification with unsupervised domain adaptation

Richard Crawley¹, Sina Amirrajab², Didier Lustermsans², Robert J. Holtackers^{3,4}, Sven Plein^{1,5}, Mitko Veta², Marcel Breeuwer², Amedeo Chiribiri¹ and Cian M. Scannell^{1,2*} 

Abstract

Quantification of myocardial scar from late gadolinium enhancement (LGE) cardiovascular magnetic resonance (CMR) images can be facilitated by automated artificial intelligence (AI)-based analysis. However, AI models are susceptible to domain shifts in which the model performance is degraded when applied to data with different characteristics than the original training data. In this study, CycleGAN models were trained to translate local hospital data to the appearance of a public LGE CMR dataset. After domain adaptation, an AI scar quantification pipeline including myocardium segmentation, scar segmentation, and computation of scar burden, previously developed on the public dataset, was evaluated on an external test set including 44 patients clinically assessed for ischemic scar. The mean \pm standard deviation Dice similarity coefficients between the manual and AI-predicted segmentations in all patients were similar to those previously reported: 0.76 ± 0.05 for myocardium and 0.75 ± 0.32 for scar, 0.41 ± 0.12 for scar in scans with pathological findings. Bland-Altman analysis showed a mean bias in scar burden percentage of -0.62% with limits of agreement from -8.4% to 7.17% . These results show the feasibility of deploying AI models, trained with public data, for LGE CMR quantification on local clinical data using unsupervised CycleGAN-based domain adaptation.

Relevance statement Our study demonstrated the possibility of using AI models trained from public databases to be applied to patient data acquired at a specific institution with different acquisition settings, without additional manual labor to obtain further training labels.

Keywords Artificial intelligence, Cardiovascular magnetic resonance, Image processing (computer assisted), Late gadolinium enhancement, Myocardium

*Correspondence:

Cian M. Scannell

cian.scannell@kcl.ac.uk

Full list of author information is available at the end of the article



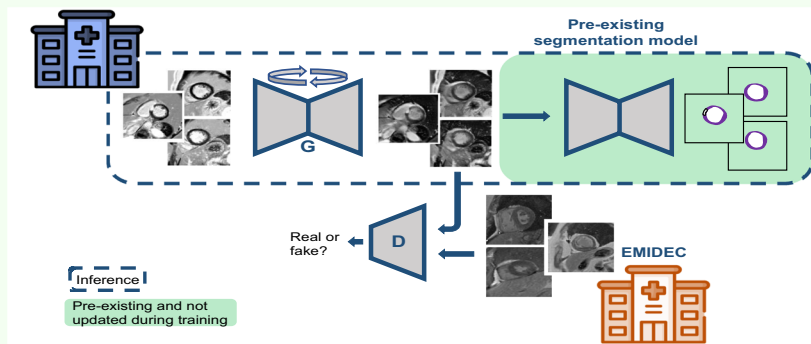
© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Graphical Abstract

Automated CMR myocardial scar quantification with unsupervised domain adaptation



- A cycle-consistent generative adversarial network (CycleGAN) can be trained to translate the appearance of local clinical late gadolinium enhancement data to match the appearance of a public dataset.
- The domain adaptation could allow the application of AI models trained on the public dataset to quantify scar on local clinical data.
- However, performance is still reduced on the new domain test set compared to the original domain warranting further work on domain adaptation and larger-scale follow-up testing.



Unsupervised CycleGAN-based domain adaptation allowed an AI pipeline for CMR scar quantification, trained on a public dataset, to be applied to local clinical data



Eur Radiol Exp (2024) Crawley R, Amirrajab S, Lustermans D et al. DOI: 10.1186/s41747-024-00497-3

Background

Late gadolinium enhancement (LGE) cardiovascular magnetic resonance (CMR) is the non-invasive reference standard to assess the presence and extent of myocardial scar [1]. Semiautomated methods for scar quantification have been previously validated as accurate [2] and reproducible [3] but the required manual contouring of the endocardial and epicardial borders remains time-consuming [4]. Therefore, a standard clinical practice still relies on image visual interpretation. Fully automated scar quantification instead could aid clinical translation and allow accurate measurements of the size and transmural-ity of myocardial infarction and could improve observer confidence.

We have previously developed an automated artificial intelligence (AI)-based CMR scar quantification pipeline [5]. The deep learning models were trained with publicly available data from a single scanner vendor as part of a competition, the “Automatic evaluation of myocardial infarction from delayed-enhancement cardiac magnetic resonance—EMIDEC” challenge [6]. Whilst such competitions are useful to promote research, models trained on selective and uniform acquisition data may not be applicable in real-world clinical applications.

It has been widely reported that acquisition-specific differences represent a “domain shift” in CMR and that this domain shift degrades the performance of deep learning models [7–9]. For LGE images, there are large differences between domains related to the acquisition parameters as well as to the type, concentration, dose, and injection protocol of the gadolinium-based contrast agent, making the application of previously trained models challenging [5]. At our institution, LGE is performed using scanners of varying field strength and from multiple vendors with a dark-blood sequence optimized to null the blood pool signal for improved scar-to-blood contrast [7]. Therefore, domain adaptation steps are required to enable clinical deployment of the EMIDEC model on our local data.

Recent work has demonstrated that cycle-consistent generative adversarial networks (CycleGANs) can be used for unpaired image-to-image translation [10], with some examples including stain normalization of histopathological slides [11]; adapting magnetic resonance imaging (MRI) images from different scanners [12, 13]; and the translation of MRI images to computed tomography [14].

The aim of this study was to develop a CycleGAN-based domain adaptation scheme that allows the use of the previously trained automatic scar quantification pipeline without additional training labels or retraining.

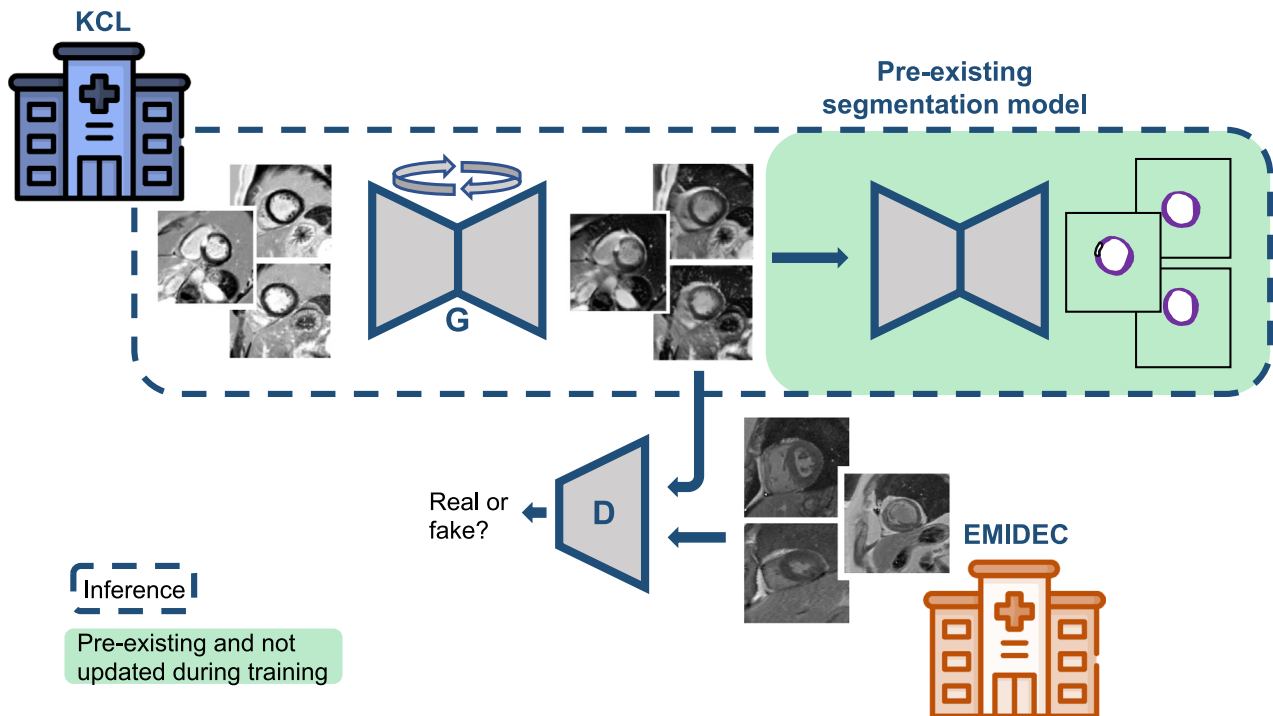


Fig. 1 Representation of the CycleGAN-based domain adaptation. Note that the CycleGAN training requires two generators and associated discriminator models (the second is not shown here, for simplicity). Hospital icons created by Freepik–Flaticon (<https://www.flaticon.com/authors/freepik>)

Methods

Datasets

The data included a development set to train the CycleGAN and an independent held-out test set to evaluate the automated scar quantification. Written informed consent was obtained from all patients for inclusion in this research. The study was approved by a UK Research Ethics Committee (reference: 15/NS/0030) and complies with the Declaration of Helsinki.

Development set

The CycleGAN training required unlabeled, unpaired images from both source and target domains. The source domain images consisted of the original EMIDEC dataset (150 patients), as previously reported [6]. The target domain images consisted of retrospectively included two-dimensional dark-blood phase-sensitive inversion-recovery LGE data from a 3-T Achieva TX scanner (Philips Healthcare, Best, the Netherlands) at King’s College London acquired at least 10 min after an intravenous injection of 0.2 mmol/kg gadobutrol (Gadovist, Bayer, Berlin, Germany), acquisition sequence as previously described [15] with further details in the Supplementary Materials. It included a convenience sample of 150 patients with 1979 slices from previous studies [16, 17].

Independent test set

For independent testing, 44 patients undergoing dark-blood LGE examinations, acquired in the same way as the development set, for clinical investigation of coronary artery disease were retrospectively enrolled. The baseline and demographic characteristics of the independent test set population are shown in Supplementary Materials. Manual labels were generated using commercially available software (cvi42, Circle CVI, Calgary, Alberta, Canada). Labeling was performed by a European Association of Cardiovascular Imaging CMR level 3 observer with > 3 years of full-time experience in CMR (R.C.). Labeling involved the manual tracing of the endocardial and epicardial boundaries followed by placing of a region of interest in the remote normal myocardium and identification of the highest signal intensity value within the whole myocardium. Areas of infarcted myocardium were identified as 5 standard deviations (SDs) above the mean signal of remote normal myocardium, as this method was found to be accurate [2] with minimal interobserver and intraobserver variability [18] for dark-blood LGE segmentation.

Automated scar quantification

The scar quantification method has been previously described [5]. It is a cascaded pipeline of three models: (1) a bounding box regression network to detect the heart; (2) a U-Net model [19] to segment the myocardium; and (3) a U-Net model to segment scar, if present.

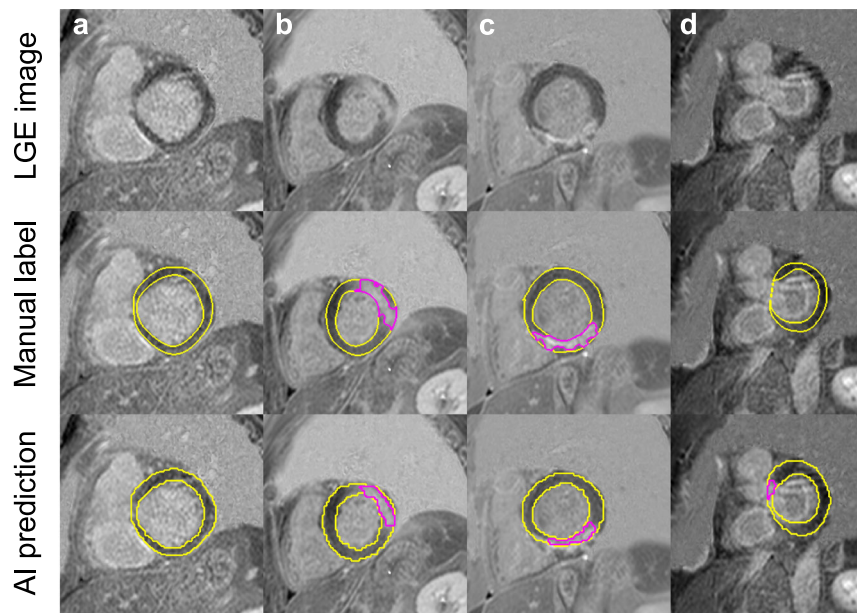


Fig. 2 A single image slice from four patients showing representative performance, comparing artificial intelligence (AI) predictions to manual myocardium (yellow) and scar (pink) contours. Column **a**: well-matched myocardium contours in a case without scar. Columns **b** and **c**: good, albeit imperfect overlap for both myocardium and scar in cases with infarction. Column **d**: false positive scar prediction by AI in a slice with visible left ventricular outflow tract

Domain adaptation

Figure 1 shows a schematic representation of the domain adaptation scheme. A CycleGAN model was trained to translate local dark-blood clinical data to appear similar to EMIDEC data so that it can be processed with the scar quantification pipeline trained on EMIDEC data. Separate CycleGAN models are used between steps 1 and 2 and between steps 2 and 3 in the scar quantification pipeline to adapt the input to the two U-Net models. CycleGAN-based domain adaptation was not used prior to the first step of the scar quantification method [5] which predicts a region of interest around the LV myocardium.

A CycleGAN used a conditional generator to map the source domain to the target domain. The generator was trained adversarially, with a discriminator attempting to distinguish real and generated images. Simultaneously, a second generator (with associated discriminator) maps data from the target domain back to the source domain. In addition to the GAN loss, cycle consistency was enforced such that an image mapped from the source to the target domain and then back (or vice versa) should match the original input, to preserve structural image information when translating one domain to the other.

The two generators used 2-D U-Net architectures [19] consisting of five resolution steps with two convolutional blocks each, with the number of convolutional filters doubling each step from 32. The training used the Adam

optimizer with $\beta_1 = 0.5, \beta_2 = 0.99$ and learning rate linearly decaying to 0 from 0.0001. The batch size was 2, affine image augmentations were used (as detailed in the Supplementary Materials), and the models were trained for 100 epochs with the best model chosen by visual assessment. The GAN loss as described by de Bel et al [11] was used with the cycle-consistency term weighted by a factor of 6 and the identity term weighted by 0.4, determined after initial experimentation. The training script is provided at <https://github.com/q-cardIA/lge-cyclegan>.

Statistical analysis

The performances of the AI scar and myocardium segmentation were assessed using the Dice similarity coefficient (DSC) which quantifies the overlap between AI and manual segmentations between 0 (no overlap) and 1 (complete overlap). Scar quantification methods were compared with a nonparametric test, *i.e.*, the Wilcoxon signed-rank test. Bland-Altman analysis of the derived scar burdens was performed to assess agreement.

Results

Representative slices from the test set are visualized in Fig. 2, comparing the myocardium and scar segmentations of the AI, with domain adaptation, *versus* the manual observer. The mean (\pm SD) DSC value between the manual and automatic segmentations was 0.76 ± 0.05

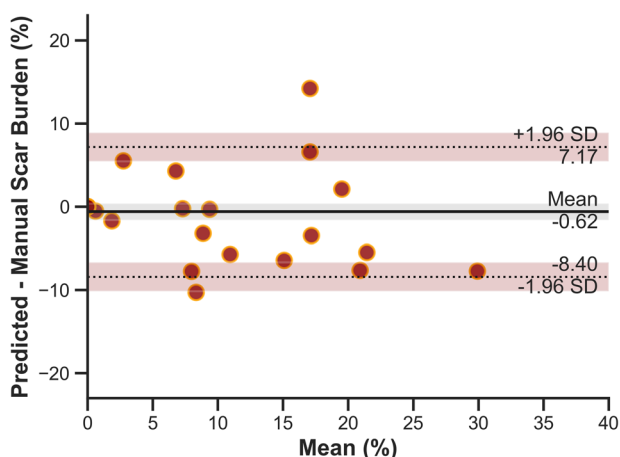


Fig. 3 Bland-Altman analysis showing a mean bias in percentage scar burden between the artificial intelligence-predicted and manual quantification of -0.62% with limits of agreement from -8.40% to 7.17% . The dotted lines represent the limits of agreement, and the shaded regions are 95% confidence intervals around each of them

for the myocardium and 0.75 ± 0.32 for the scar. The overall scar segmentation DSC was biased by the perfect overlap in patients without scar and the scar DSC in the subset of patients that were positive for scar was 0.41 ± 0.12 . There was no statistically significant difference between the manual scar analysis and the automated AI scar burden quantification ($p = 0.170$). Figure 3 shows the Bland-Altman analysis comparing the scar burden. Total run time was lower than 3 min per patient including data loading, processing, and all model inference.

Discussion

In this study, we developed an unsupervised domain adaptation scheme using CycleGAN models to translate the image appearance of a single vendor publicly available LGE CMR data to match local clinical data acquired using a different scanner vendor with different acquisition parameters. We demonstrated the possibility of using AI models trained from public databases to be applied to patient data acquired at a specific institution with different acquisition settings. Of note, this approach did not require additional manual labor to obtain further training labels and could help to exploit the growing amounts of publicly available data.

Previous studies proposing automated myocardial scar quantification have trained and tested their models using a single cohort from a single study at one field strength with similar acquisition parameters and no external validation [20]. The performance reported in our study is slightly lower than those previously published [20], but this is expected given our unsupervised domain adaptation approach with the external test dataset. The mean (\pm SD) DSC value between the manual and automatic segmentations was 0.76 ± 0.05 for the myocardium and

0.75 ± 0.32 for the scar. The higher variability of scar DSC is due to the relatively small size of scar regions to be segmented which means that even small misalignments of the AI segmentation can lead to low DSC values [21] and because patients with no scar correctly predicted as having no scar achieve a perfect DSC of 1, so there is a wide range of scar DSC values. Nevertheless, the automated scar burden showed reasonable agreement with manual quantification (Fig. 3), and this agreement would be expected to increase if the pipeline were to be re-trained with data from the target domain.

Due to its time-intensive nature, myocardial scar quantification is not routinely performed in clinical practice. Furthermore, the lack of availability of labeled data has resulted in scar quantification rarely being studied with automated AI algorithms. An automated algorithm could lead to increased routine use, leading to a more reproducible and accurate assessment tool. In addition, automated AI-based quantification of scars could be combined with automated quantitative myocardial perfusion assessment [22].

In conclusion, this study demonstrates that unsupervised domain adaptation using a CycleGAN model may facilitate the clinical deployment of AI models across centers. Our application for LGE imaging in CMR could allow automated scar quantification of dark-blood local clinical data using a model previously trained on public data. It is acknowledged that this is a small retrospective evaluation, only considering ischemic scar and the clinical utility of the scar quantification tool will need to be assessed prospectively at a larger scale with scars of different etiology.

Abbreviations

AI	Artificial intelligence
CMR	Cardiovascular magnetic resonance
CycleGAN	Cycle-consistent generative adversarial network
DSC	Dice similarity coefficient
EMIDEC	Automatic evaluation of myocardial infarction from delayed-enhancement cardiac magnetic resonance
LGE	Late gadolinium enhancement
MRI	Magnetic resonance imaging
SD	Standard deviation

Supplementary information

The online version contains supplementary material available at <https://doi.org/10.1186/s41747-024-00497-3>.

Additional file 1: Supplementary Table 1: Patient demographics (N = 44), values are n (%) or mean \pm SD. Left ventricular end-diastolic volume is indexed to body surface area calculation. **Supplementary Table 2:** The parameters used for the data augmentation in the bounding box training. U(a, b) denotes that the parameter value was randomly sampled from a uniform distribution on the interval [a, b]. Translation and scaling were applied independently in x and y. Parameters for translation and scaling are given as a proportion of the image size.

Acknowledgements

Large language models were not used for this manuscript.

Authors contributions

RC: conceptualization, methodology, data curation validation, writing—original draft, writing—review and editing. SA: conceptualization, methodology, writing—review and editing. Supervision. DL: conceptualization, methodology, Software, writing—review and editing. RJH: conceptualization, data curation, writing—review and editing. SP: conceptualization, methodology, writing—review and editing, supervision. MV: conceptualization, methodology, writing—review and editing, Supervision. MB: conceptualization, methodology, writing—review and editing, supervision. AC: conceptualization, methodology, writing—review and editing, supervision. CMS: conceptualization, methodology, software, writing—original draft, writing—review and editing, supervision, project administration.

Funding

The authors acknowledge financial support from: the Wellcome Trust [WT 222678/Z/16/Z] and the Wellcome/EPSCRC Center for Medical Engineering [WT 203148/Z/16/Z]. For the purpose of open access, the corresponding author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

Data availability

The code for model training is provided at <https://github.com/q-cardia/lge-cyclegan>. The EMIDEC challenge data is available at <https://emidec.com/>. Local clinical data is not made publicly available due to institutional restrictions.

Declarations

Ethics approval and consent to participate

Written informed consent was obtained from all patients for inclusion in this research. The study was approved by a UK Health Research Authority Research Ethics Committee (ref: 15/NS/0030) and complies with the Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹School of Biomedical Engineering & Imaging Sciences, King's College London, London, UK. ²Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, the Netherlands. ³Cardiovascular Research Institute Maastricht (CARIM), Maastricht University, Maastricht, the Netherlands. ⁴Department of Radiology and Nuclear Medicine, Maastricht University Medical Center, Maastricht, the Netherlands. ⁵Leeds Institute of Cardiovascular and Metabolic Medicine, University of Leeds, Leeds, UK.

Received: 3 January 2024 Accepted: 15 July 2024

Published online: 14 August 2024

References

- Kramer CM, Barkhausen J, Bucciarelli-Ducci C et al (2020) Standardized cardiovascular magnetic resonance imaging (CMR) protocols: 2020 update. *J Cardiovasc Magn Reson* 22:1–18. <https://doi.org/10.1186/s12968-020-00607-1>
- Nies HMJM, Gommers S, Bijvoet GP et al (2023) Histopathological validation of semi-automated myocardial scar quantification techniques for dark-blood late gadolinium enhancement magnetic resonance imaging. *Eur Heart J Cardiovasc Imaging* 24:364–372. <https://doi.org/10.1093/ehjci/jeac107>
- Flett AS, Hasleton J, Cook C et al (2011) Evaluation of techniques for the quantification of myocardial scar of differing etiology using cardiac magnetic resonance. *JACC Cardiovasc Imaging* 4:150–156. <https://doi.org/10.1016/J.JCMG.2010.11.015>

- Papetti DM, Van Abeelen K, Davies R et al (2023) An accurate and time-efficient deep learning-based system for automated segmentation and reporting of cardiac magnetic resonance-detected ischemic scar. *Comput Methods Programs Biomed* 229:107321. <https://doi.org/10.1016/J.CMPB.2022.107321>
- Lustermans, DRPRM, Amirrajab S, Veta M et al (2022) Optimized automated cardiac MR scar quantification with GAN-based data augmentation. *Comput Methods Programs Biomed* 226:107116. <https://doi.org/10.1016/J.CMPB.2022.107116>
- Lalande A, Chen Z, Decourselle T et al (2020) Emidec: a database usable for the automatic evaluation of myocardial infarction from delayed-enhancement cardiac MRI. *Data* 5:4. <https://doi.org/10.3390/data5040089>
- Campello VM, Gkontra P, Izquierdo C et al (2021) Multi-centre, multi-vendor and multi-disease cardiac segmentation: the M and Ms challenge. *IEEE Trans Med Imaging* 1–1. <https://doi.org/10.1109/TMI.2021.3090082>
- Yu AC, Mohajer B, Eng J (2022) External validation of deep learning algorithms for radiologic diagnosis: a systematic review. *Radiol Artif Intell* 4:3. <https://doi.org/10.1148/Ryai.210064>
- Cohen JP, Hashir M, Brooks R, Bertrand H (2020) On the limits of cross-domain generalization in automated X-ray prediction. *Proc Mach Learn Res* 121:136–155
- Zhu J-Y, Park T, Isola P, Efros AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. In: 2017 IEEE International Conference on Computer Vision (ICCV). pp 2242–2251
- de Bel T, Bokhorst J-M, van der Laak J, Litjens G (2021) Residual cyclegan for robust domain transformation of histopathological tissue slides. *Med Image Anal* 70:102004. <https://doi.org/10.1016/j.media.2021.102004>
- Yan W, Huang L, Xia L et al (2020) MRI manufacturer shift and adaptation: increasing the generalizability of deep learning segmentation for MR images acquired with different scanners. *Radiol Artif Intell* 2:e190195. <https://doi.org/10.1148/ryai.2020190195>
- Jaspers TJM, Martens B, Crawley R et al (2024) Deep learning synthesis of white-blood from dark-blood late gadolinium enhancement cardiac magnetic resonance. *Invest Radiol*. <https://doi.org/10.1097/RLI.0000000000001086>
- Kearney V, Ziemer BP, Perry A et al (2020) Attention-aware discrimination for MR-to-CT image translation using cycle-consistent generative adversarial networks. *Radiol Artif Intell* 2:e190027. <https://doi.org/10.1148/ryai.2020190027>
- Holtackers RJ, Chiribiri A, Schneider T et al (2017) Dark-blood late gadolinium enhancement without additional magnetization preparation. *J Cardiovasc Magn Reson* 19:64. <https://doi.org/10.1186/s12968-017-0372-4>
- Holtackers RJ, Van De Heyning CM, Nazir MS et al (2019) Clinical value of dark-blood late gadolinium enhancement cardiovascular magnetic resonance without additional magnetization preparation. *J Cardiovasc Magn Reson* 21:44. <https://doi.org/10.1186/s12968-019-0556-1>
- Lim RP, Kachel S, Villa ADM et al (2022) CardiSort: a convolutional neural network for cross-vendor automated sorting of cardiac MR images. *Eur Radiol* 32:5907–5920. <https://doi.org/10.1007/s00330-022-08724-4>
- Jada L, Holtackers RJ, Martens B et al (2024) Quantification of myocardial scar of different etiology using dark- and bright-blood late gadolinium enhancement cardiovascular magnetic resonance. *Sci Rep* 14:1–9. <https://doi.org/10.1038/s41598-024-52058-8>
- Ronneberger O, Fischer P, Brox T (2015) U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF (eds) *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015*. Springer International Publishing, Cham, pp 234–241
- Ghanbari F, Joyce T, Lorenzoni V et al (2023) AI cardiac MRI scar analysis aids prediction of major arrhythmic events in the multicenter DERIVATE registry. *Radiology* 307:e222239. <https://doi.org/10.1148/radiol.222239>
- Maier-Hein L, Reinke A, Godau P et al (2024) Metrics reloaded: recommendations for image analysis validation. *Nat Method* 21:195–212. <https://doi.org/10.1038/s41592-023-02151-z>
- Scannell CM, Veta M, Villa ADM et al (2020) Deep-learning-based pre-processing for quantitative myocardial perfusion MRI. *J Magn Reson Imaging* 51:1689–1696. <https://doi.org/10.1002/jmri.26983>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.