

# Physics-informed self-supervised deep learning reconstruction for accelerated first-pass perfusion cardiac MRI

Elena Martín-González<sup>1[0000-0002-5922-4960]</sup>, Ebrahim Alskaf<sup>2[0000-0003-4007-6854]</sup>, Amedeo Chiribiri<sup>2[0000-0003-3394-4289]</sup>, Pablo Casaseca-de-la-Higuera<sup>1[0000-0003-1565-0842]</sup>, Carlos Alberola-López<sup>1[0000-0003-3684-0055]</sup>, Rita G. Nunes<sup>3[0000-0001-7425-5717]</sup>, and Teresa Correia<sup>2,4[0000-0002-1606-9550]</sup>

<sup>1</sup> Laboratorio de Procesado de Imagen, ETSI Telecomunicación, Universidad de Valladolid, Valladolid, Spain  
 emargon@lpi.tel.uva.es

<sup>2</sup> School of Biomedical Engineering and Imaging Sciences, King's College London, London, United Kingdom

<sup>3</sup> Institute for Systems and Robotics and Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

<sup>4</sup> Centre for Marine Sciences - CCMAR, Faro, Portugal

**Abstract.** First-pass perfusion cardiac magnetic resonance (FPP-CMR) is becoming an essential non-invasive imaging method for detecting deficits of myocardial blood flow, allowing the assessment of coronary heart disease. Nevertheless, acquisitions suffer from relatively low spatial resolution and limited heart coverage. Compressed sensing (CS) methods have been proposed to accelerate FPP-CMR and achieve higher spatial resolution. However, the long reconstruction times have limited the widespread clinical use of CS in FPP-CMR. Deep learning techniques based on supervised learning have emerged as alternatives for speeding up reconstructions. However, these approaches require fully sampled data for training, which is not possible to obtain, particularly high-resolution FPP-CMR images. Here, we propose a physics-informed self-supervised deep learning FPP-CMR reconstruction approach for accelerating FPP-CMR scans and hence facilitate high spatial resolution imaging. The proposed method provides high-quality FPP-CMR images from 10x undersampled data without using fully-sampled reference data.

**Keywords:** Deep learning reconstruction · Model-based reconstruction · Quantitative perfusion cardiac MRI

## 1 Introduction

Coronary artery disease (CAD) is the occlusion of the coronary arteries usually caused by atherosclerosis, which causes abnormalities in blood flow to the heart. Non-invasive imaging techniques that are widely used clinically for the evaluation of CAD are single photon emission computerized tomography (SPECT)

and positron emission tomography (PET), but the reference for non-invasive myocardial perfusion quantification is PET [8]. However, the clinical value of first-pass perfusion cardiac magnetic resonance (FPP-CMR) has been shown in comparison to these techniques [6,7,8,20], having emerged as an alternative way of detecting blood flow anomalies without the use of potentially harmful ionising radiation. In addition, FPP-CMR has other advantages, such as higher spatial resolution, wider availability and lower scan cost compared to PET.

FPP-CMR time frames must be acquired in real-time to capture the rapid passage of a contrast agent bolus through the heart, and hence, the spatial resolution and coverage of the heart is compromised. Thus, undersampled reconstruction methods have been proposed to accelerate FPP-CMR acquisitions as a means to improve spatial resolution and heart coverage [14,16,21]. However, these methods can lead to long reconstruction times. In this work, we aim to speed up reconstructions and obtain the contrast-enhanced dynamic image series from undersampled FPP-CMR using deep learning (DL). Then, these images will be used to generate quantitative perfusion maps using a tracer kinetic model [4,9,11]. DL techniques have already been used in magnetic resonance image (MRI) reconstruction. Work has been reported on knee [2,13,22], brain [2,5,13,22] and cardiac [10,19] MRI, using both supervised [2,5,13] and self-supervised learning [15,22]. Occasionally, the network is unrolled to mimic a compressed sensing (CS) iterative reconstruction problem, giving rise to a cascade of convolutional neural networks (CNNs) [2,13,19]. The problem with supervised learning techniques is the need to have fully sampled reference images to train the network, which are not available in FPP-CMR, particularly at high spatial resolutions.

Even though the field of MRI reconstruction with DL is currently an active area, to our knowledge, self-supervised DL techniques have not been applied to FPP-CMR reconstruction. In this work, a SElf-Supervised aCcelerated REconsTruction (SECRET) DL framework for FPP-CMR is proposed to directly reconstruct contrast-enhanced dynamic image series from undersampled (k,t)-space data.

## 2 Methods

For completeness, a conventional FPP-CMR CS reconstruction will be described. We will also describe our proposed method, SECRET, as well as the Model Based Deep Learning Architecture for Inverse Problems (MoDL) [2], which will be used for comparison.

### 2.1 Conventional FPP-CMR reconstruction

CS methods can be used to reconstruct dynamic images from undersampled data. For example, FPP-CMR images  $\mathbf{s}$  can be obtained from undersampled data  $\mathbf{d}_u$  using CS by solving the following optimisation problem:

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s}} \{ \|\mathbf{d}_u - \mathbf{E}\mathbf{s}\|_2^2 + \lambda_1 \|\nabla_s \mathbf{s}\|_1 + \lambda_2 \|\nabla_t \mathbf{s}\|_1 \} \quad (1)$$

where  $\mathbf{E} = \mathbf{A}\mathcal{F}$ ,  $\mathbf{A}$  is the (k,t)-space sampling trajectory,  $\mathcal{F}$  is the Fourier transform,  $\lambda_1$  and  $\lambda_2$  are regularization parameters and  $\nabla_s$  and  $\nabla_t$  are the finite differences operators along the spatial and temporal dimensions, respectively.

## 2.2 Supervised learning reconstruction: MoDL

MoDL combines the power of DL with model-based approaches [2]. It uses a CNN as a denoiser and applies it as a regulariser to solve the optimisation problem given by:

$$\mathbf{s}_{k+1} = \arg \min_{\mathbf{s}} \|\mathbf{d}_u - \mathbf{E}\mathbf{s}\|_2^2 + \lambda \|\mathbf{s} - \mathbf{z}_k\|_2^2 \quad (2)$$

$$\mathbf{s}_{k+1} = (\mathbf{E}^H \mathbf{E} + \lambda \mathcal{I})^{-1} (\mathbf{E}^H \mathbf{d}_u + \lambda \mathbf{z}_k) \quad (3)$$

where  $k$  denotes the  $k$ -th iteration and  $\mathbf{z}_k$  is the denoised version of  $\mathbf{s}_k$ , obtained through a CNN network. MoDL requires supervised learning to optimise the denoiser network. The data consistency layer is immediate by conjugate gradient blocks, but as the input is  $\mathbf{z}_k$  and the output is  $\mathbf{s}_{k+1}$ , which, in turn, generates a  $\mathbf{z}_{k+1}$ , this requires iterating until convergence. The iterative algorithm is unrolled for a fixed number of iterations,  $K$ , in which the weights or parameters to be optimised are shared.

The MoDL method has the zero-filled reconstruction, the coil sensitivities and the subsampling mask as inputs, but it also needs the fully sampled images — which are hardly available for the case of FPP-CMR at high spatial resolution— for training. The loss is defined as the mean square error between  $s_K$  and the desired image  $t$ :  $C = \sum_{i=1}^{N_{samples}} \|\mathbf{s}_K(i) - \mathbf{t}(i)\|^2$ , where  $\mathbf{t}(i)$  is the  $i$ -th target image.

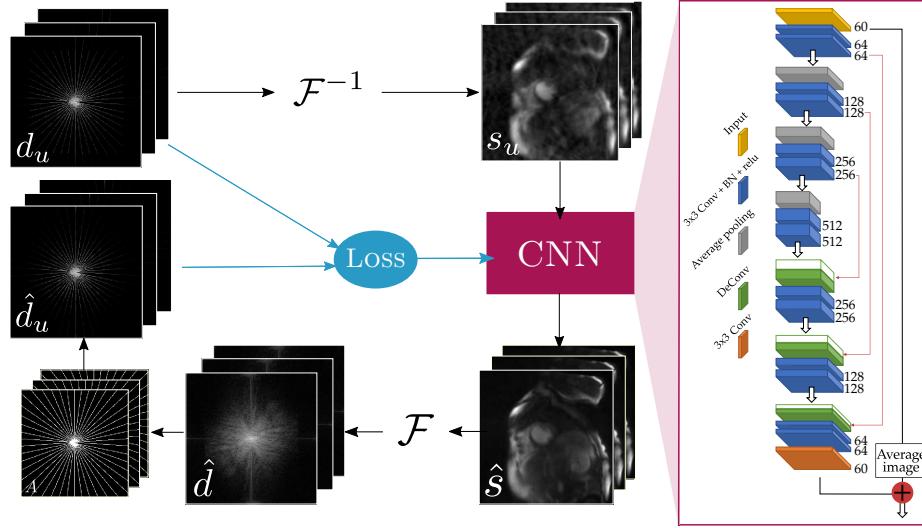
## 2.3 SECRET reconstruction

The proposed SECRET method directly reconstructs contrast-enhanced dynamic images from the undersampled (k,t)-space data. Considering only the undersampled (k,t)-space data when enforcing data consistency, we can train networks without the need for fully sampled images, simply by making use of the physical models in the reconstruction [15]. This framework can be formulated as follows:

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \|\mathbf{d}_u - \mathbf{A}\mathcal{F}C(\mathbf{s}_u | \boldsymbol{\theta})\|_2 \quad (4)$$

where  $C(\mathbf{s}_u | \boldsymbol{\theta})$  is the output of a CNN, with  $\boldsymbol{\theta}$  the parameter vector to be optimised. Figure 1 shows the steps necessary for training our proposed SECRET method for FPP-CMR. First, undersampled (k,t)-space data  $\mathbf{d}_u$  is transformed to the image domain, obtaining  $\mathbf{s}_u$ . Then,  $\mathbf{s}_u$  enters the CNN to provide the reconstructed contrast-enhanced dynamic images  $\hat{\mathbf{s}}$ . These images are then transformed back to (k,t)-space  $\hat{\mathbf{d}}$  and subsampling masks are applied, thus obtaining the undersampled version  $\hat{\mathbf{d}}_u$ . Finally, the loss is computed with  $\hat{\mathbf{d}}_u$  and the input  $\mathbf{d}_u$ , to guide the training phase.

The CNN is based on the well-known U-Net [18], widely used in medical imaging. Skip connections are included to maintain information from previous layers, as well as to avoid the problem of vanishing gradients during backpropagation. At the end of the CNN, residual learning has been appended as in [15], adding the average image of the input  $s_u$ .



**Fig. 1.** Flow chart illustrating the proposed SECRET method for FPP-CMR. Blue lines represent steps that only take place during training. The inputs of the framework are the undersampled  $(k,t)$ -space data  $\mathbf{d}_u$  and the  $(k,t)$ -sampling masks  $A$ , resulting in the reconstructed contrast-enhanced dynamic images  $\hat{\mathbf{s}}$  as output, and  $\hat{\mathbf{d}}_u$  if required.

## 2.4 Dataset

Rest and stress FPP-CMR acquisitions were performed in 21 patients using a single-bolus injection of 0.05 mmol/kg Gadobutrol (Gadovist; Bayer, Germany) and a 1.5T CMR scanner (MAGNETOM Aera, Siemens Healthineers, Erlangen, Germany) with an 18-channel chest-coil and a 32-channel spine coil. A free-breathing FLASH perfusion dual-sequence [11] was used to acquire a low-resolution image with low T1-sensitivity for estimating the arterial input function and three short-axis slices (basal, mid and apical) for high resolution myocardial perfusion imaging using the following parameters: FOV =  $340 \times 308$  mm $^2$ , in-plane resolution =  $2.2 \times 2.2$  mm $^2$ , slice thickness = 10mm, TR/TE = 2.1/1ms, flip angle =  $8^\circ$ , parallel imaging acceleration factor 3, saturation recovery time = 100 ms, total scan duration = 60s, contrast agent relaxivity = 5.0L/mmol s. Undersampled datasets were generated for  $3\times$ ,  $6\times$  and  $10\times$  acceleration factors, following a radial  $(k,t)$ -sampling trajectory.

**Preprocessing** A first step to ensure that all data had the same size, both spatially and temporally, prior to being fed to the CNN, consisted of resizing the DICOM images to obtain a spatial resolution of  $2 \times 2 \text{ mm}^2$ , padding the k-space to obtain an image size of  $256 \times 256$  pixels, and interpolating each slice to a fixed number of frames (60 frames). A final step included intensity normalisation so that all contrast-enhanced dynamic image series present intensities between 0 and 1, without losing the contrast variation between frames. In addition, image pre-registration was also carried out to correct for respiratory motion.

**Image quality metrics** Image quality was assessed in terms of peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM) and normalized root mean square error (NRMSE) between the reference images and reconstructions obtained with the SECRET, MoDL and CS (10x only) methods.

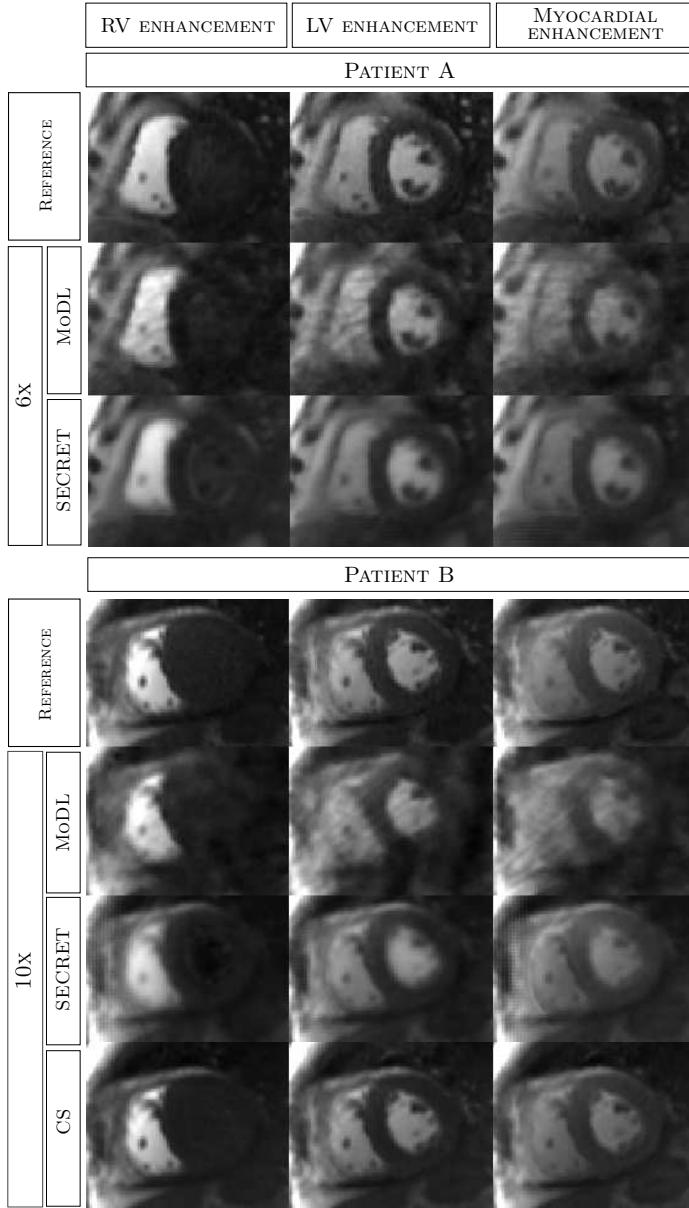
## 2.5 Implementation details

Patients were randomly split into training, validation and test subsets (60%, 16% and 24%, respectively). Each slice is fed into the SECRET framework so that the time frames are stacked in depth, creating a multi-channel image. The proposed method is implemented in Python with Tensorflow [1] and Keras [3], and it took about half an hour of training using the Adam optimizer [12] with a learning rate of  $10^{-4}$  consuming about 3 GB of GPU memory for 100 epochs on one Intel® Core™ i7-4790 CPU @ 3.60GHz with 16 GB RAM and one NVIDIA GeForce RTX 2080 Ti GPU. The MoDL training for  $K=1$  and 100 epochs took one hour and a half and the MoDL training for  $K=10$  and 200 epochs took forty-five hours using the same hardware. Note that after training the SECRET method, it provides a reconstruction of a complete contrast-enhanced dynamic image series in less than a second.

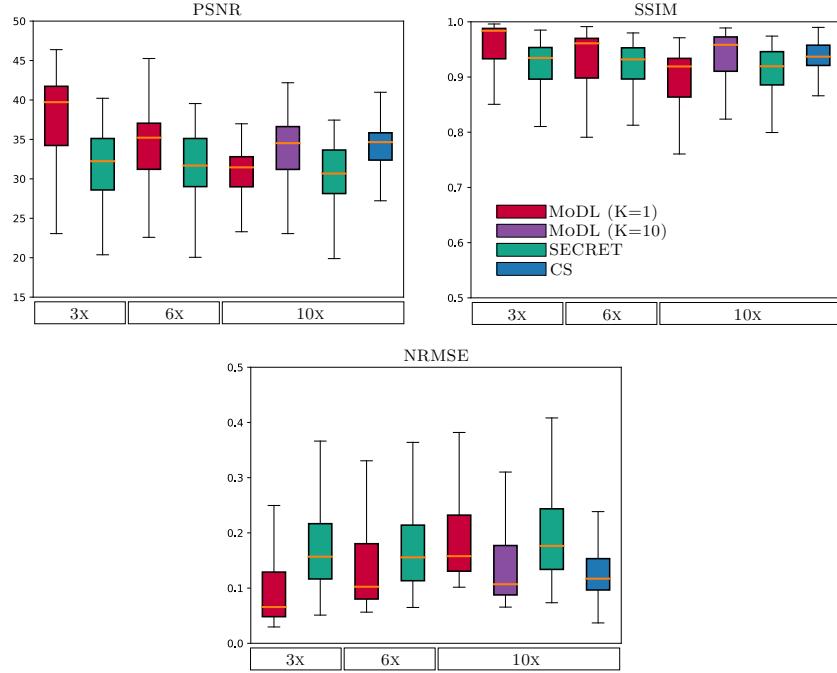
## 3 Results and discussion

Figure 2 shows the SECRET reconstructions obtained for two representative patients from  $6\times$  and  $10\times$  undersampled ( $k,t$ )-space data together with the reference and MoDL ( $K=1$ ) reconstructions. CS reconstruction is also shown for  $10\times$ . Three different time frames are shown, corresponding to right ventricle (RV), left ventricle (LV) and myocardial enhancement. Although the SECRET reconstructions are slightly blurred, due to residual learning from the average image of the CNN input (which is blurred due to residual motion), it can be seen that they have better quality than the images obtained with MoDL trained in the same amount of time. Moreover, SECRET images maintain the variability of contrast that exists between frames in addition to not losing the structure of the heart.

Figure 3 shows results of the FPP-CMR reconstructions in terms of PSNR, SSIM and NRMSE. While the performance of MoDL becomes noticeably worse as the acceleration rate increases, SECRET maintains good image quality even at high acceleration rates. For the 10x accelerated reconstructions, the **median**



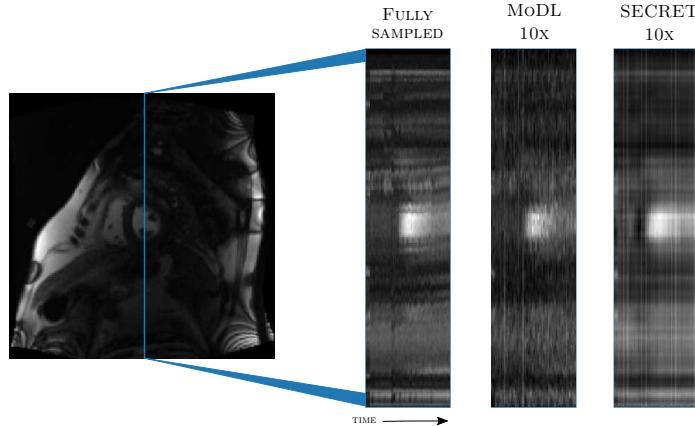
**Fig. 2.** SECRET and MoDL ( $K=1$ ) reconstructions obtained from  $6\times$  and  $10\times$  undersampled FPP-CMR data for two representative subjects. The reference images are displayed for comparison, in addition to CS reconstruction for  $10\times$ . The right ventricle (RV), left ventricle (LV) and myocardial enhancement time frames are shown for one short axis slice.



**Fig. 3.** PSNR, SSIM and NRMSE between the reference images and the reconstructions obtained with SECRET and MoDL methods, for 3 $\times$ , 6 $\times$  and 10 $\times$  acceleration factors, for all patients in the test dataset.

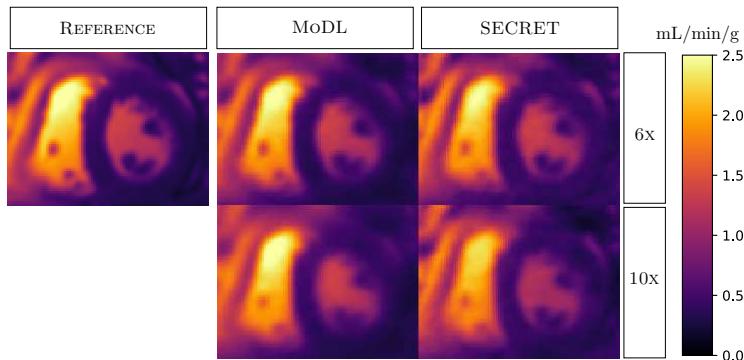
(interquartile range): PSNR was **34.66** (3.47), **31.46** (3.81), **34.52** (5.43), **30.67** (5.52); SSIM was **0.94** (0.04), **0.92** (0.07), **0.96** (0.06), **0.92** (0.06); NRMSE was **0.12** (0.06), **0.16** (0.10), **0.11** (0.09), **0.17** (0.11) for CS, MoDL (K=1), MoDL (K=10) and SECRET methods, respectively. The image quality metrics indicate that SECRET images maintain a more stable agreement with the reference as the acceleration factor is increased than MoDL images, which deteriorate with higher acceleration. CS and MoDL (K=10) show the best agreement with the reference, but reconstructions take  $\sim$ 87.08s and  $\sim$ 1.99s, respectively, whereas MoDL (K=1) takes  $\sim$ 0.21s and SECRET only 0.15s.

Figure 4 shows a 1D projection of the dynamic images through time, for a given slice. Note that although the images have been pre-registered, there is still some residual motion. SECRET does not include any explicit regularisation term, however, due to the residual learning performed by the network all reconstructions provided by the framework are inherently corrected. Such good PSNR, SSIM and NRMSE values obtained when the reference images are affected by little respiratory motion, would certainly improve if some regularisation were added. This would enable even higher acceleration rates. Regularisation schemes will thus be investigated in a future study.



**Fig. 4.** Representative image profile across the heart demonstrating that the SECRET framework improves consistency across time frames.

Quantitative parameter maps were estimated from the FPP-CMR reconstructions, showing the potential of the technique for an objective and operator-independent analysis of myocardial perfusion. Figure 5 displays the contrast transfer coefficient ( $K^{\text{Trans}}$ ) map estimated from fully sampled,  $6\times$  and  $10\times$  undersampled patient data using the MoDL and SECRET methods, through the Patlak model [17]. The image quality of the quantitative maps obtained from the SECRET reconstruction at accelerations  $6\times$  and  $10\times$  is comparable to the reference images, showing less blurring than MoDL maps.



**Fig. 5.** Quantitative maps ( $K^{\text{Trans}}$ ) obtained from  $6\times$  and  $10\times$  undersampled data using MoDL and the SECRET methods. The reference image is displayed for comparison.

## 4 Conclusion

A physics-informed self-supervised deep learning reconstruction framework for accelerating FPP-CMR scans has been described. The proposed SECRET method provides FPP-CMR reconstructions directly from the undersampled  $(k,t)$ -space data and does not require fully sampled reference data. Compared with state-of-the-art approaches, the SECRET method maintains good quality reconstructions for higher acceleration rates, with low training times and very fast reconstruction times. The proposed SECRET method shows promising results, with the potential for improvement coupled with explicit regularization, which will be explored in future work.

## References

1. Abadi, M., et al.: TensorFlow: Large-scale machine learning on heterogeneous systems (2015), <https://www.tensorflow.org/>, software available from tensorflow.org
2. Aggarwal, H.K., Mani, M.P., Jacob, M.: Modl: Model-based deep learning architecture for inverse problems. *IEEE transactions on medical imaging* **38**(2), 394–405 (2018)
3. Chollet, F., et al.: Keras. <https://keras.io> (2015)
4. Correia, T., Schneider, T., Chiribiri, A.: Model-based reconstruction for highly accelerated first-pass perfusion cardiac mri. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 514–522. Springer (2019)
5. Do, W.J., Seo, S., Han, Y., Ye, J.C., Choi, S.H., Park, S.H.: Reconstruction of multicontrast mr images through deep learning. *Medical physics* **47**(3), 983–997 (2020)
6. Foley, J.R.J.: Cardiovascular Magnetic Resonance Imaging for the Investigation of Ischaemic Heart Disease. Ph.D. thesis, University of Leeds (2018)
7. Hendel, R.C., Friedrich, M.G., Schulz-Menger, J., Zemmrich, C., Bengel, F., Berman, D.S., Camici, P.G., Flamm, S.D., Le Guludec, D., Kim, R., et al.: Cmr first-pass perfusion for suspected inducible myocardial ischemia. *JACC: Cardiovascular imaging* **9**(11), 1338–1348 (2016)
8. Heo, R., Nakazato, R., Kalra, D., Min, J.K.: Noninvasive imaging in coronary artery disease. In: Seminars in nuclear medicine. vol. 44, pp. 398–409. Elsevier (2014)
9. Hsu, L.Y., Jacobs, M., Benovoy, M., Ta, A.D., Conn, H.M., Winkler, S., Greve, A.M., Chen, M.Y., Shanbhag, S.M., Bandettini, W.P., et al.: Diagnostic performance of fully automated pixel-wise quantitative myocardial perfusion imaging by cardiovascular magnetic resonance. *JACC: Cardiovascular Imaging* **11**(5), 697–707 (2018)
10. Huang, Q., Yang, D., Wu, P., Qu, H., Yi, J., Metaxas, D.: Mri reconstruction via cascaded channel-wise attention network. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). pp. 1622–1626. IEEE (2019)
11. Kellman, P., Hansen, M.S., Nielles-Vallespin, S., Nickander, J., Themudo, R., Ugander, M., Xue, H.: Myocardial perfusion cardiovascular magnetic resonance: optimized dual sequence and reconstruction for quantification. *Journal of Cardiovascular Magnetic Resonance* **19**(1), 1–14 (2017)

12. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
13. Kocanaogullari, D., Eksioglu, E.M.: Deep learning for mri reconstruction using a novel projection based cascaded network. In: 2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP). pp. 1–6. IEEE (2019)
14. Lingala, S.G., Hu, Y., DiBella, E., Jacob, M.: Accelerated dynamic mri exploiting sparsity and low-rank structure: kt slr. *IEEE transactions on medical imaging* **30**(5), 1042–1054 (2011)
15. Liu, F., Kijowski, R., El Fakhri, G., Feng, L.: Magnetic resonance parameter mapping using model-guided self-supervised deep learning. *Magnetic Resonance in Medicine* (2021)
16. Otazo, R., Kim, D., Axel, L., Sodickson, D.K.: Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion mri. *Magnetic resonance in medicine* **64**(3), 767–776 (2010)
17. Patlak, C.S., Blasberg, R.G., Fenstermacher, J.D.: Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. *Journal of Cerebral Blood Flow & Metabolism* **3**(1), 1–7 (1983)
18. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. pp. 234–241. Springer (2015)
19. Schlemper, J., Caballero, J., Hajnal, J.V., Price, A.N., Rueckert, D.: A deep cascade of convolutional neural networks for dynamic mr image reconstruction. *IEEE transactions on Medical Imaging* **37**(2), 491–503 (2017)
20. Schwitter, J., Wacker, C.M., Wilke, N., Al-Saadi, N., Sauer, E., Huettel, K., Schönberg, S.O., Luchner, A., Strohm, O., Ahlstrom, H., et al.: Mr-impact ii: Magnetic resonance imaging for myocardial perfusion assessment in coronary artery disease trial: perfusion-cardiac magnetic resonance vs. single-photon emission computed tomography for the detection of coronary artery disease: a comparative multicentre, multivendor trial. *European heart journal* **34**(10), 775–781 (2013)
21. Vitanis, V., Manka, R., Giese, D., Pedersen, H., Plein, S., Boesiger, P., Kozerke, S.: High resolution three-dimensional cardiac perfusion imaging using compartment-based k-t principal component analysis. *Magnetic resonance in medicine* **65**(2), 575–587 (2011)
22. Yaman, B., Hosseini, S.A.H., Moeller, S., Ellermann, J., Uğurbil, K., Akçakaya, M.: Self-supervised learning of physics-guided reconstruction neural networks without fully sampled reference data. *Magnetic resonance in medicine* **84**(6), 3172–3191 (2020)