

Deep variational network for rapid 4D flow MRI reconstruction

Valery Vishnevskiy^{1*}, Jonas Walheim^{1*}, and Sebastian Kozerke¹

¹Institute for Biomedical Engineering, University and ETH Zurich, Zurich, Switzerland

*Equal contribution

April 22, 2020

Abstract

Phase-contrast (PC) magnetic resonance imaging provides time-resolved quantification of blood flow dynamics that can aid clinical diagnosis. Long in vivo scan times due to repeated 3D volume sampling over cardiac phases and breathing cycles necessitate accelerated imaging techniques that leverage data correlations. Standard compressed sensing reconstruction methods require tuning of hyperparameters and are computationally expensive, which diminishes potential reduction of examination times. We propose an efficient model-based deep neural reconstruction network and evaluate its performance on clinical aortic flow data. The network is shown to reconstruct undersampled 4D flow MRI data in under a minute on the standard consumer hardware. Remarkably, the relatively low amounts of tunable parameters allowed the network to be trained on images from 11 reference scans while generalizing well to retrospective and prospective undersampled data for various acceleration factors and anatomies.

Introduction

4D flow MRI provides spatio-temporally resolved quantification of blood flow and offers great potential for the assessment of cardiovascular disease, e.g. aortic valve stenosis, atherosclerosis, or vessel wall remodelling¹. However, clinical adaptation of the method has been hampered by long exam times.

Many efforts have been dedicated to accelerate flow acquisition by exploiting redundancies in the data. Partial Fourier imaging² has been used for moderate acceleration³, but the underlying assumption of a slowly varying phase has been shown to be incorrect for 4D flow MRI⁴. Parallel imaging (PI)⁵, which exploits the spatially varying sensitivity of receiver elements in the coil array has become a standard for accelerated imaging, but undersampling rates are limited by noise amplification⁶. The advent of compressed sensing (CS)⁷ has enabled acceleration of 4D flow MRI by acquiring only a subset of k-space data and exploiting prior information about data regularities during reconstruction^{8,9,10,11,12,13}, with typical acceleration factors ranging from 5¹⁰ up to 27¹³. In particular, the locally low rank (LLR) regularized reconstruction¹⁴ has been a successful technique, which iteratively balances the data fidelity cost and the singular norm of a patch matrix stacked over cardiac phases (see Methods for details). However, iterative reconstruction methods as used in CS increase reconstruction times considerably, implying that evaluation of 4D flow MRI data will typically happen when the subject has already been moved out of the scanner.

In recent years, deep neural networks have gained increasing popularity in MR image reconstruction. In the training stage, the neural network learns abstract features from a set of scans. After training, newly acquired data are reconstructed with very little computational effort by inference with the learned weights. This reduction in reconstruction times can facilitate the use of accelerated imaging methods in clinical practice. Moreover, reconstruction results can be superior to traditional CS methods^{15,16}. Some approaches discard concepts of iterative image reconstruction altogether, e.g. by learning end-to-end mappings from k-space to image space¹⁷. As a downside, such networks usually require abundant amounts of high quality training data which is not available for high-dimensional flow MRI. Model-based neural reconstruction networks can also be designed to replicate the behaviour of an iterative reconstruction by interlacing nonlinear

convolutional filters with an operation that enforces closeness of the current image estimate to the acquired data^{15,18,16,19}, similar to the data fidelity step in an iterative shrinkage-thresholding algorithm²⁰. A recent study²¹ showed that neural network architectures which incorporate such an operation generalize better to different undersampling rates. In contrast, generic architectures which are solely based on convolutional layers can even lead to deteriorated image quality when the undersampling factor is increased, although one would expect the reconstruction result to improve when more information is available. The adversarial approach for training MR reconstruction networks^{22,23} is usually aimed at improving perceptual reconstruction quality, such as image sharpness. Typically, this is achieved at the expense of reconstruction nRMSE (normalized root-mean-square error)²⁴, which is critical for flow quantification.

In this work, an approach based on the idea of deep variational neural networks¹⁵ is implemented for rapid 4D flow reconstruction, which is referred to as FlowVN hereafter. For comparison, the 3D variational network (VN) architecture as presented in¹⁵ was adapted for 4D flow data by using 3D filter banks operating on subsets of $xyzt$ dimensional data, yielding the model that we refer to as HamVN. The FlowVN network architecture replicates 10 steps of an iterative image reconstruction, while allowing for learnable spatio-temporal filter kernels, activation functions, and regularization weights in each iteration. It is demonstrated that based on training performed with retrospectively undersampled data of healthy subjects, FlowVN can accurately reconstruct pathological flow in a stenotic aorta in 21 seconds. Moreover, an imaging study with healthy subjects demonstrates good agreement of reconstructions from prospective undersampling with reference measurement.

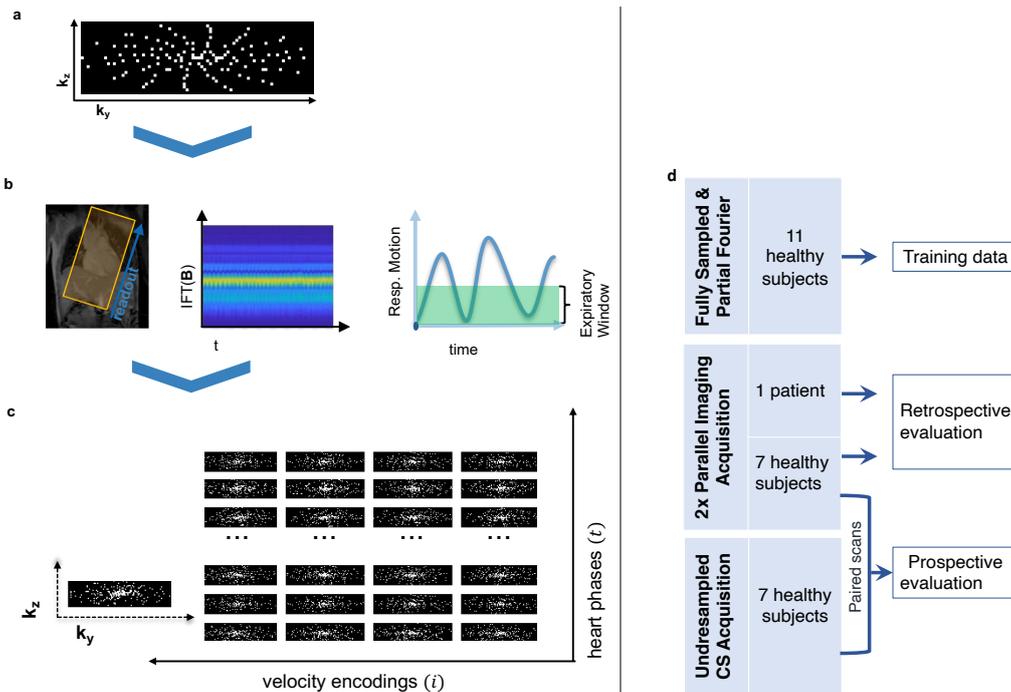


Figure 1: **Breathing resolved 4D flow data acquisition.** **a**, Data are sampled using a Cartesian pseudo-radial tiny Golden angle sampling pattern. **b**, Time-resolved data are binned to end expiration using a combination of principal component analysis, low-pass filtering, and coil-clustering. **c**, Acquired data are sorted according to heart phase and velocity encoding. **d**, Datasets used during training and evaluation. To conduct prospective evaluation, 7 healthy volunteers underwent accelerated CS and reference PI acquisitions during the same scan session.

4D flow MRI reconstruction with FlowVN

The FlowVN architecture improves HamVN in the following ways: (i) linear activations are used instead of radial basis functions (RBF), (ii) the network is conditioned on the sampling rate, (iii) exponential weighting of intermediate layers is used as regularization, (iv) real and imaginary parts of the signal are filtered by shared weights, (v) momentum is considered during gradient

Table 1: **Model complexities and typical reconstruction time for 4D flow reconstruction.** Typical reconstruction time for 4-point velocity encoded data compressed to 5 virtual coils and reconstructed on $113 \times 113 \times 25$ grid. CS-LLR was executed on a 6-core Intel CPU, FlowVN and HamVN were implemented in Tensorflow and evaluated on a NVIDIA Titan RTX.

Method	Recon. Time	# of Param.
CS-LLR	10 min 24 s	2
HamVN	89 s	62,742
FlowVN	21 s	63,583

descent (GD) unrolling, and, (vi) the data term allows tunable activation functions. The network is trained for a wide range of acceleration factors by allowing acceleration dependent weighting of data consistency and filtering steps.

As illustrated in Fig. 1a, for each velocity encoding direction, the k-space data is acquired using a Cartesian Golden angle sampling strategy, yielding variable density undersampling patterns in k-space. The signal of a total of 28 physical coils is compressed into 5 virtual coils via clustering²⁵. The samples are then sorted into respiratory bins and data in the end-expiratory bin is used for reconstruction.

A deep variational network can be seen as a differentiable sequence of an unrolled numerical optimization scheme. To enable learning, such sequence is then relaxed by allowing tunable filter weights and activation functions. As described in Methods, we unroll $K=10$ steps of a gradient descent with momentum governed by a scalar $\alpha^{(k)}$:

$$\mathbf{S}^{(k+1)} \leftarrow \alpha^{(k+1)} \mathbf{S}^{(k)} + \mathbf{G}^{(k)}, \quad (1)$$

$$\mathbf{P}^{(k+1)} \leftarrow \mathbf{P}^{(k)} - \mathbf{S}^{(k+1)}. \quad (2)$$

At each k -th layer $\mathbf{P}^{(k)}$ the current complex-valued spatio-temporal image estimate is represented, while $\mathbf{S}^{(k)}$ maintains a running average of update steps. The update step $\mathbf{G}^{(k)}$ consists of the data consistency and regularization terms (see Methods and Supplementary Algorithm 1 for details), that are weighted according to the sampling rate $\bar{M} = 1/R$ (R is the acceleration factor) via tunable activation functions $\varphi_{\text{ud}}^{(k)}(\bar{M})$ and $\varphi_{\text{ur}}^{(k)}(\bar{M})$, respectively. The data consistency term modulates the k-space data residual via an activation function and maps them back to the image space via a conjugate imaging operator. The regularization term at each layer contains 3D filters grouped into 4 banks, where each bank performs convolutions in 3 dedicated dimensions, namely xyz , xyt , xzt and yzt , therefore avoiding costly 4D convolutions. To avoid overfitting, we assume shared filters and activation functions that operate on real and imaginary components of the image. Note that both data and regularization terms do not assume correlations between real and imaginary parts of the signal, as highlighted in Fig. 2b.

The image estimate $\mathbf{P}^{(K)}(\mathbf{B}, \Theta)$ of the final layer can be then seen as a function of the k-space samples \mathbf{B} and network parameters Θ . To tune the network parameters Θ we minimize the layer-wise exponentially weighted ℓ_1 image reconstruction loss:

$$\min_{\Theta} \mathbb{E}_{\{\mathbf{B}, \mathbf{P}^*\} \sim \mathcal{T}} \sum_{k=1}^K e^{-\tau(K-k)} \|\mathbf{P}^{(k)}(\mathbf{B}; \Theta) - \mathbf{P}^*\|_1, \quad (3)$$

over the retrospectively undersampled training dataset \mathcal{T} , where \mathbf{P}^* is the ground truth image. Layer weighting is controlled by parameter $\tau \geq 0$: when $\tau \approx 0$, the reconstruction error is penalized equally across layers, therefore gradients of network parameters have lower variance during stochastic optimization, yielding faster convergence. On the contrary, when $\tau \rightarrow +\infty$, only reconstruction at the final layer $\mathbf{P}^{(K)}$ is minimized, which improves fitting accuracy on the training data. It is worth mentioning that τ controls the trade-off between training reconstruction residual and network regularity. Similarly to Landweber iterations^{26,15} and deep supervision²⁷, such implicit regularization penalizes irregular representations at intermediate layers and favors networks that can provide fast reconstruction. We propose to initialize τ with zero and then gradually increase it according to the training schedule (see Methods).

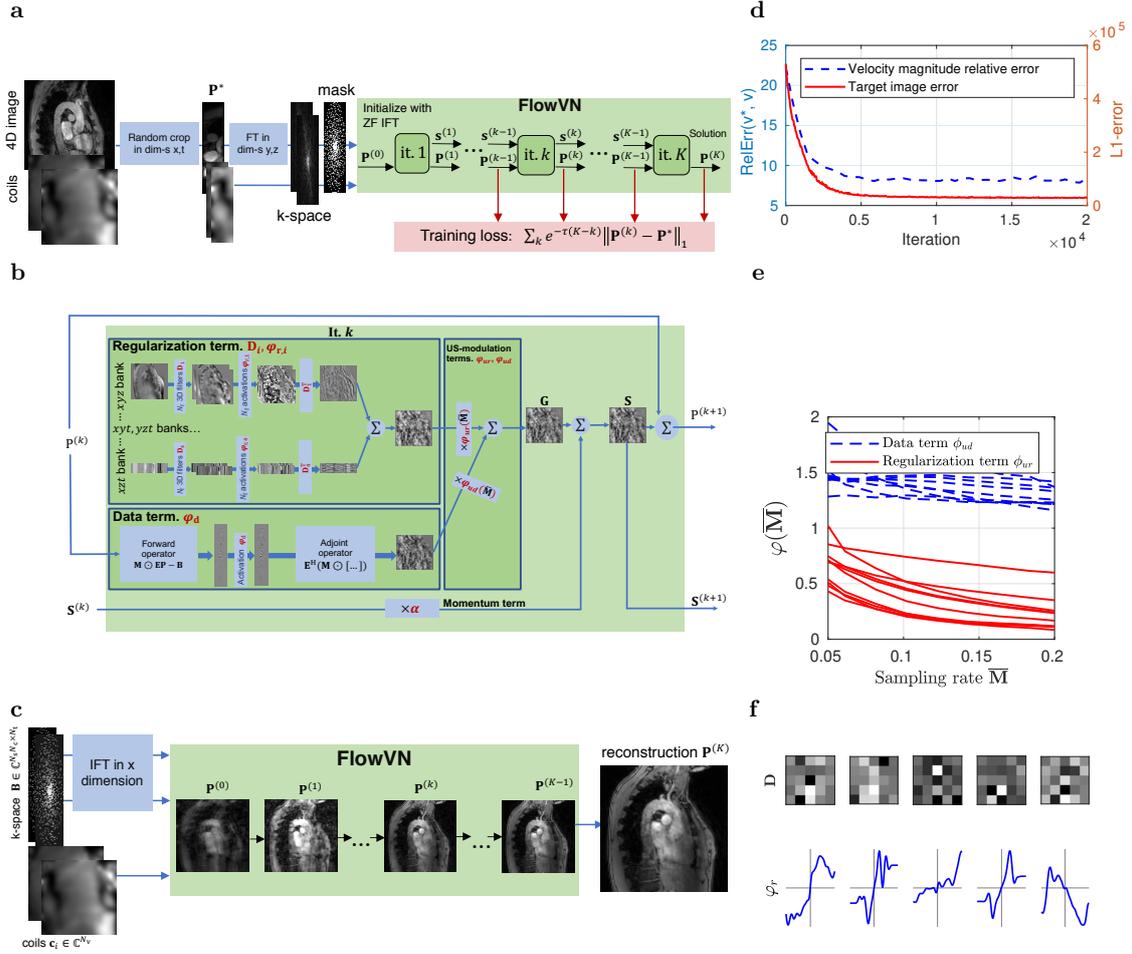


Figure 2: **FlowVN architecture and training.** **a**, Structure of FlowVN and its training strategy which uses a reduced field-of-view data. **b**, Single unrolled iteration block consisting of convolutional filtering, data fidelity, undersampling modulation and gradient descent momentum terms. Tunable parameters are highlighted in red. **c**, FlowVN at inference time yields 4D image reconstruction. **d**, Target training image error and velocity magnitude error in the aorta evaluated using training data. **e**, Data and gradient term weighting functions shown for each of 10 layers. **f**, Exemplary slices of 3D xyz filters and their corresponding activation functions at layer 5.

To demonstrate the validity of our approach, we note that the extracted velocity magnitude error in the aorta decreases simultaneously with the target reconstruction error during training, as shown in Fig. 2d, therefore indicating that the target ℓ_1 image error is a valid training surrogate. It can be seen from Fig. 2e that the regularization term is suppressed for lower acceleration factors R (higher sampling rate \bar{M}). A subset of learned FlowVN parameters Θ is shown in Fig. 2f illustrating that learned convolutions perform direction-dependent filtering.

Retrospective and Prospective Evaluation

Reconstructed image magnitudes (for a single velocity encoding component), estimated velocity magnitudes and their errors of a healthy volunteer data for acceleration factor $R=14$ are illustrated in Fig. 3 for retrospectively undersampled data. Compared to CS-LLR and HamVN, the proposed FlowVN provides better reconstruction accuracy in terms of image magnitude and velocities. Scatter plot and correlation analysis further suggest that the velocity magnitude image estimated via FlowVN is in better agreement with ground truth. As shown in Supplementary Table 2 these observations extend to other acceleration factors R (6–22) as tested on 7 healthy volunteers.

Fig. 4 indicates that FlowVN can accurately reconstruct the jet at the inlet section of the aorta for a patient with a pathological aortic valve.

The prospective undersampling acquisition results are reported in Fig. 5a,b: peak velocities and

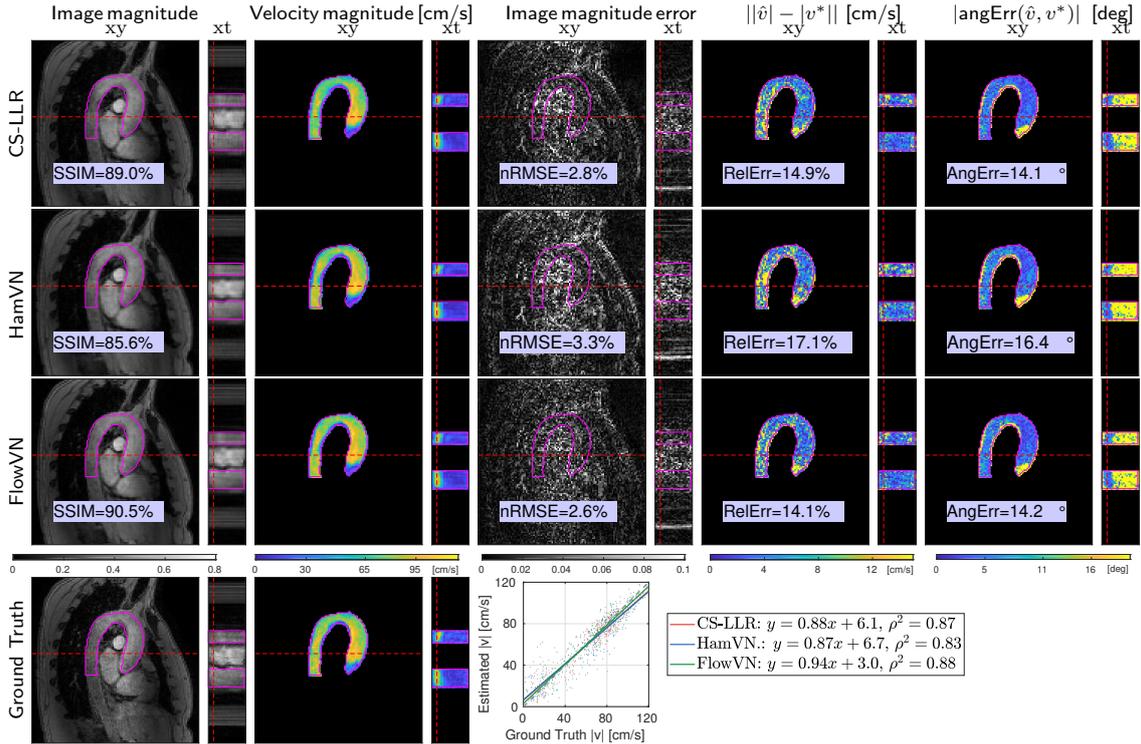


Figure 3: **Reconstruction results on retrospectively undersampled data.** Image magnitudes and estimated 4D velocity magnitude maps on retrospectively $14\times$ undersampled data from a healthy volunteer. Corresponding slice locations are illustrated with red dashed lines, indicating crosssection of the aorta and systolic peak. Scatter plot of velocity magnitude over manually segmented aorta (contour shown in magenta) is given together with correlation analysis ($y = ax + b$).

peak flow estimated using CS-LLR and FlowVN are in good agreement with PI reconstruction, while HamVN systematically underestimates velocity magnitudes. Moreover, correlation analysis in Fig. 5d reveals high correlation between CS-LLR and FlowVN velocity estimates. In contrast, HamVN shows systematic velocity underestimation, compared to CS-LLR.

The exemplary reconstruction time for typical 4-point velocity encoded images reported in Table 1 shows that the proposed FlowVN is 30 times faster than CS-LLR reconstruction.

Discussion

Practical learning-based image reconstruction can be traced back to dictionary learning methods^{28,29}, where prior information is learned from image patches and then used as a sparsity-inducing regularizer for iterative reconstruction. Such an approach yields orders of magnitude longer reconstruction times, compared to modern deep learning approaches. A straightforward application of deep artificial neural networks has been suggested for learning reconstruction as a regression from the k-space¹⁷ or zero-filled reconstructions³⁰ directly into the image space. Although tempting, such an approach might be unjustified, because the k-space and zero-filling artefacts have global dependence on image intensities. The advent of effective automatic differentiation systems^{31,32} revitalized the idea of unrolling³³ and relaxing numerical schemes that can solve the original reconstruction problem. Following this approach, a number of deep neural network architectures were proposed^{18,34,15,35}, that disentangle image acquisition and image prior models. Unrolling gradient descent reconstruction with tunable filters and activation functions yields the HamVN architecture proposed by Hammernik et al.¹⁵. One advantage of VNs is that, compared to other deep architectures, they employ relatively limited number of free parameters to tune, therefore they are less susceptible to overfitting.

In this work we have further developed the VN architecture^{15,36,37} to accommodate high performance undersampled 4D flow reconstruction with limited training. Namely, we avoid exponential model complexity growth by avoiding 4D convolutions and by using separable 3D convolutions

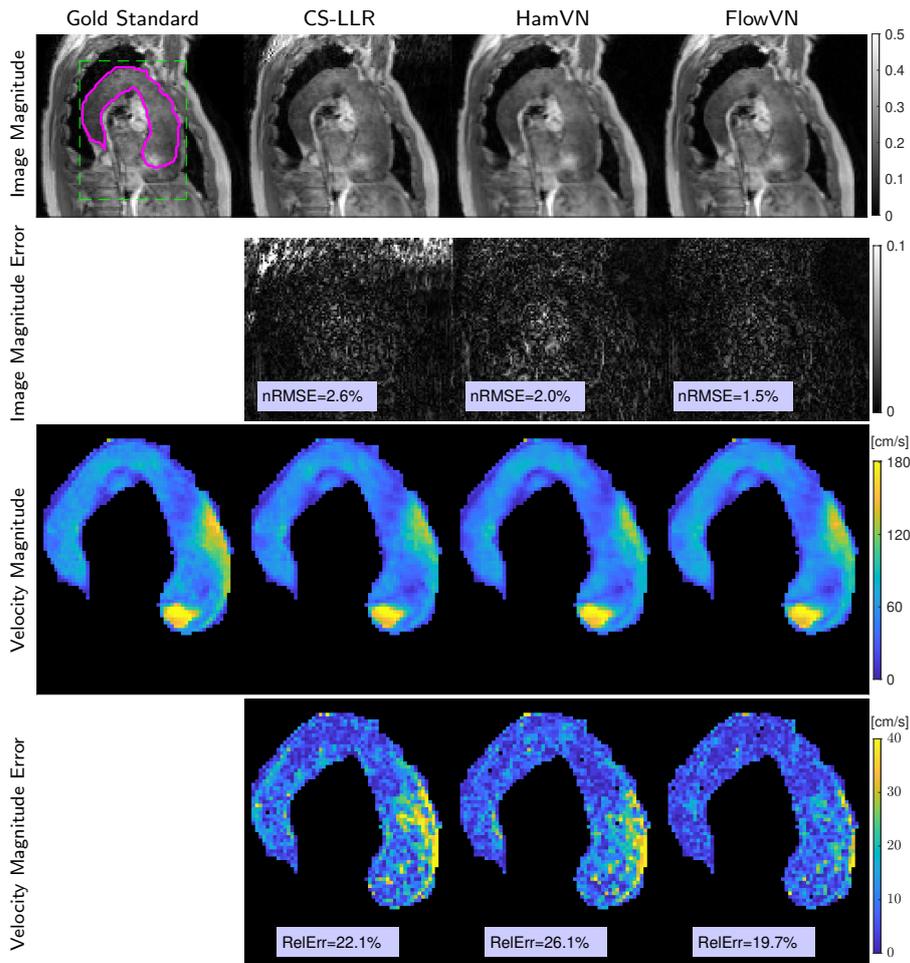


Figure 4: **Retrospective reconstruction of the data from patient with abnormal flow pattern.** Reconstruction results of $10\times$ retrospectively undersampled patient data shown at systolic peak flow. Manual aorta segmentation and field-of-view are shown with magenta and green dashed lines, respectively.

that are shared for real and imaginary parts of the image. Furthermore, in contrast to the original HamVN¹⁵, we train our FlowVN for a wide range of undersampling factors by allowing the regularization term to depend on them. As illustrated in Fig. 2e, regularization scaled by φ_{ur} decreases as more samples are available, while the data term φ_{ud} stays constant for most of the layers. Such conditioning allows network training on a larger variety of artefacts and as it is necessary in practice, since for a given fixed acquisition time, the precise value of the undersampling factor is not known a priori and depends on breathing and cardiac motion patterns. We hypothesize that the wide range of acceleration factors which were used simultaneously to train the FlowVN provided a diverse collection of aliasing artefacts and enabled robust learning on a remarkably limited training set of 11 subjects. The exponential weighting of the layer-wise reconstruction loss (3) further regularized FlowVN parameters by penalizing the nonlinear behaviour presented in HamVN reconstructions. Supplemental Fig. 6 and Table 3 provide quantification of reconstruction accuracy effects attributed to the modifications proposed with FlowVN. In particular, modifications to the network architecture result in a model that can better adapt to data and yield higher accuracy for retrospectively undersampled experiments, while the proposed exponential weighting of the training loss improves accuracy of the prospective evaluation, which indicates better generalization ability. It is worth noting, that FlowVN has only 1% more tunable parameters compared to HamVN (c.f. Table 1), while improving reconstruction nRMSE by 23% (averaged over acceleration factors as given in Supplemental Fig. 6). We note that 4D flow MRI greatly benefits from using coil information during reconstruction (see Supplementary Table 3). Accordingly, comparison with single-coil reconstruction networks¹⁸ has limited benefit.

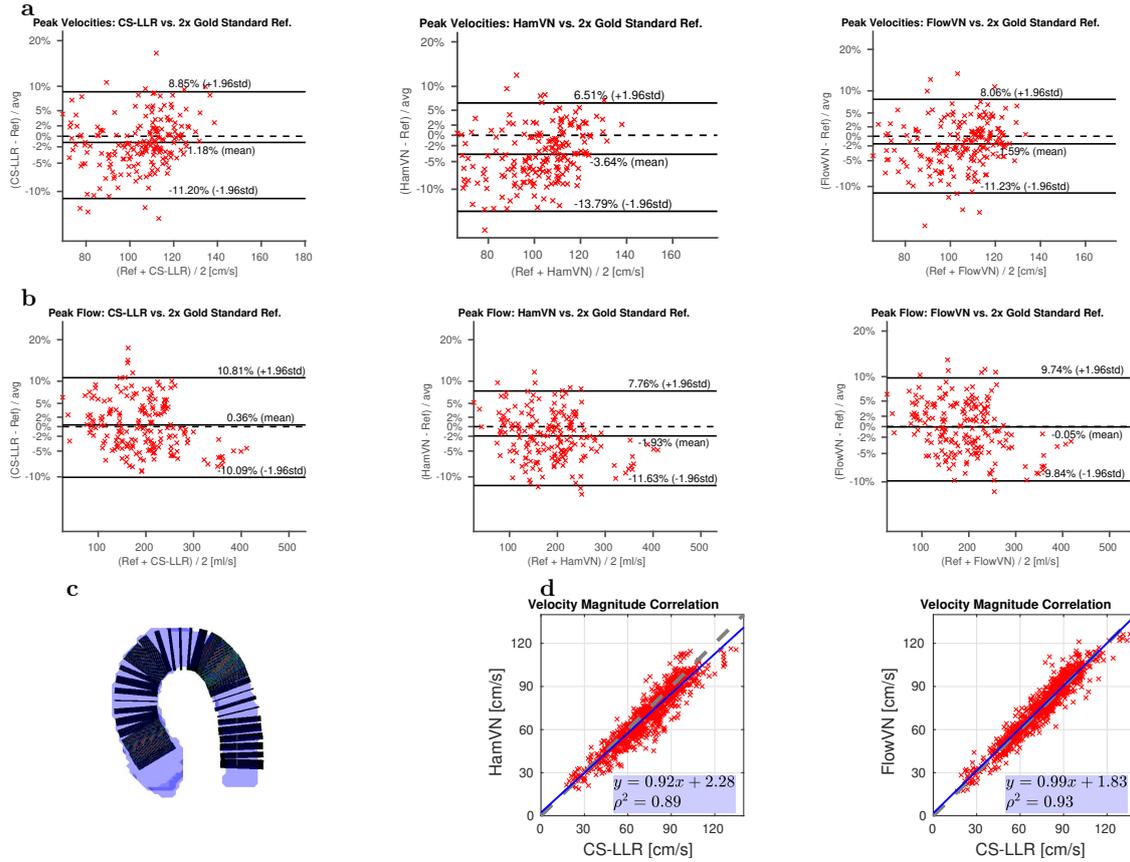


Figure 5: **Quantitative flow evaluation of reconstruction methods on the prospectively undersampled data ($12.4 \leq R \leq 13.8$) from 7 healthy volunteers.** **a,b**, Bland-Altman analysis of peak velocities and peak flow assessed over manually segmented aorta slices as illustrated in **c**. Black dashed bars indicate zero difference between corresponding measurements and solid lines show mean and standard deviation values. **d**, Correlation analysis of aortic velocity magnitude estimates by learned VN architectures and CS-LLR reconstruction. The blue line shows the linear regression fit and grey dashed line corresponds to $y = x$.

The proposed FlowVN is a learning-based approach for reconstructing undersampled 4D flow MRI data in under a minute. For fixed reconstruction accuracy, FlowVN enables higher acceleration factors (12% improvement compared to CS-LLR image nRMSE at $R=16$) and does not introduce significant bias of peak flow estimates. The proposed reconstruction is 30 times faster than state-of-the-art CS-LLR and 4.2 times faster than HamVN, due to using linear activation functions rather than RBFs, which requires computation of pairwise distances between control knots and image intensities. It is worth noting that FlowVN demonstrates high generalization ability, being able to preserve patient pathologies that were not present in the training data.

Methods

Compressed Sensing 4D Flow Reconstruction

PC MRI encodes flow velocity $\mathbf{v}(\mathbf{r}, t) \in \mathbb{R}^3$ at spatial location \mathbf{r} during cardiac phase t ($1 \leq t \leq N_t$) according to the following equation:

$$\rho_i(\mathbf{r}, t) = \rho_0(\mathbf{r}, t) \exp\left(j\pi \frac{(\Phi \mathbf{v}(\mathbf{r}, t))_i}{v_{\text{enc}}}\right), \quad (4)$$

where v_{enc} is the velocity corresponding to a phase of $\pm\pi$, $y_i, i = 0, \dots, 3$ are the encoded velocity vector components. The four-point velocity encoding matrix is given as

$$\Phi = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (5)$$

Therefore, flow velocity \mathbf{v} can be calculated from the phase difference of reconstructed PC images ρ_i .

Let $\rho_{it} \in \mathbb{C}^{N_v}$ be a discretized image on a $N_x \times N_y \times N_z = N_v$ grid corresponding to a cardiac phase t and velocity encoding i . Assuming Cartesian sampling on a regular $N_1 \times N_2 \times N_3 = N_s$ grid, the Fourier transform $\mathbf{F} \in \mathbb{C}^{N_s \times N_v}$, and N_c coil sensitivity maps $\mathbf{W}_k \stackrel{\text{def}}{=} \text{diag}(\mathbf{c}_k) \in \mathbb{C}^{N_v \times N_v}$ define the spatial encoding operator $\mathbf{E} \in \mathbb{C}^{N_s N_c \times N_v}$:

$$\mathbf{E}\boldsymbol{\rho} \stackrel{\text{def}}{=} \left[(\mathbf{F}\mathbf{W}_1\boldsymbol{\rho})^\top, \dots, (\mathbf{F}\mathbf{W}_{N_c}\boldsymbol{\rho})^\top \right]^\top \in \mathbb{C}^{N_s N_c}. \quad (6)$$

Considering a single velocity-encoded image sequence, let $\mathbf{P} \in \mathbb{C}^{N_v \times N_t}$ and $\mathbf{B} \in \mathbb{C}^{N_s N_c \times N_t}$ be stacked column-vectors of signals $\boldsymbol{\rho}$ and zero-filled k-space samples respectively, while $\mathbf{M} \in \{0, 1\}^{N_s N_c \times N_t}$ defines the undersampling mask. Iterative image reconstruction methods seek for a maximum a posteriori (MAP) solution defined by the following optimization problem:

$$\hat{\mathbf{P}}_{\text{MAP}} = \underset{\mathbf{P}}{\text{argmin}} \frac{1}{2} \|\mathbf{M} \odot (\mathbf{E}\mathbf{P} - \mathbf{B})\|_{\mathbb{F}}^2 + \mathcal{R}(\mathbf{P}), \quad (7)$$

where the regularization term \mathcal{R} enforces prior assumptions about image regularities. Herein we consider the local low-rank (LLR) regularization¹⁴ to leverage image correlations among cardiac phases:

$$\mathcal{R}_{\text{LLR}}(\mathbf{P}) = \lambda_{\text{LLR}} \sum_{i \leq N_{\text{ptch}}} \|\mathbf{T}_i \mathbf{P}\|_*, \quad (8)$$

where $\mathbf{T}_i \in \{0, 1\}^{p^3 \times N_v}$ is the corresponding $p \times p \times p$ patch extraction operator, yielding N_{ptch} patches, and $\|\cdot\|_*$ is the nuclear norm. For LLR regularization the optimization problem (7) is convex and can be efficiently solved using operator splitting techniques such as the fast iterative shrinkage-thresholding algorithm (FISTA)²⁰.

FlowVN Training

We employ a $K=10$ layer VN and perform $5 \cdot 10^4$ iterations of the ADAM algorithm (learning rate 10^{-3} , $\beta_1=0.85$, $\beta_2=0.98$, batch size of 3) for training, during which we continually adjust $\tau = i_{\text{opt}} \cdot 10^{-3}$ with i_{opt} being the iteration number. On every layer, each 3D filter bank contains $N_f=8$ filters of size $n_c=5$ voxels. Activation functions $\varphi\{\cdot\}$ are parametrized by $N_{\text{knts}}=91$ control knots with spacing $\omega=0.17$:

$$\varphi\{h\} = (1 - h/\omega + \lfloor h/\omega \rfloor) \phi_{\lfloor h/\omega \rfloor} + (h/\omega - \lfloor h/\omega \rfloor) \phi_{\lfloor h/\omega \rfloor + 1}, \quad (9)$$

with gradients provided by the following formulas:

$$\begin{aligned} \frac{\partial \varphi}{\partial \phi_i} \{h\} &= \mathbf{1}_{i \leq h \leq i+1} \cdot (1 - h + \lfloor h \rfloor) + \\ &\quad \mathbf{1}_{i-1 \leq h \leq i} \cdot (h - \lfloor h \rfloor), \\ \frac{\partial \varphi}{\partial h} \{h\} &= \phi_{\lfloor h \rfloor + 1} - \phi_{\lfloor h \rfloor}. \end{aligned} \quad (10)$$

The acquired zero-filled k-space \mathbf{B} with undersampling mask \mathbf{M} was normalized by $\frac{\|\mathbf{M}\|_1}{\|\mathbf{B}\|_F}$.

To enable backpropagation to be carried out with limited GPU memory, we employ spatio-temporal equivariance of the convolution and exploit the fact that k-space is fully sampled in the readout dimension k_x for Cartesian acquisitions. Therefore, to draw a training sample, we perform random cropping of width w_x and w_t in dimensions x and t respectively and simulate Fourier encoding in $k_y k_z$ dimensions as illustrated in Fig. 2(a). The network was implemented using the Tensorflow framework³². Fully-sampled and partial Fourier acquisition data from 11 healthy volunteers was used during training.

In Vivo Data Acquisition

As illustrated in Fig. 1, we used 11 subject for network training, and 7 healthy subjects and 1 patient for evaluation. All in-vivo work was performed upon written informed consent of the subjects and according to local ethics regulations.

Training datasets comprised 4D flow data measured in the aorta of 11 healthy subjects, 9 of them fully sampled and 2 acquired with partial Fourier³⁸ (factor 0.75×0.75).

For evaluation, data in the ascending aorta of 7 healthy subjects were acquired on a 3T Philips Ingenia system (Philips Healthcare, Best, the Netherlands) using a Cartesian 4-point referenced phase-contrast gradient-echo sequence with an encoding velocity $v_{\text{enc}} = 150 \text{ cm/s}$, a spatial resolution of $2.5 \times 2.5 \times 2.5 \text{ mm}^3$, $T_E = 3.3 \text{ ms}$, $T_R = 4.9 \text{ ms}$, 25 cardiac phases and flip angle $= 8^\circ$. Exams for each of the 7 healthy subjects comprised a standard navigator-gated 2-fold accelerated parallel imaging⁵ exam for reference, and a CS acquisition with an acceleration factor of $R=12.4 - 13.8$, using Cartesian pseudo-radial golden angle sampling pattern³⁹ and data driven-respiratory motion detection, as in¹¹. Only data in expiration were kept for reconstruction as shown in Fig. 1.

To evaluate reconstruction accuracy on pathological anatomy, 4D flow data was acquired in a single patient with dilation of the ascending aorta, combined aortic stenosis and regurgitation due to a bicuspid aortic valve on a 3T Philips Ingenia system (Philips Healthcare, Best, the Netherlands) using a navigator-gated 2-fold accelerated parallel imaging⁵ scan.

A receiver coil with 28 channels was used for acquisition which were reduced to 5 channels using coil compression⁴⁰. Coil sensitivity maps were estimated with ESPIRiT⁴¹. Concomitant field correction was applied to the signal phase according to Bernstein et al.⁴² and eddy currents were corrected for with a third-order polynomial model fitted to stationary tissue^{43,44}.

Evaluation

We compared the proposed FlowVN to the state-of-the-art compressed sensing LLR-regularized (8) reconstruction¹⁴ and the variational network by Hammernik et al.¹⁵ which we refer to as HamVN. The LLR implementation from the Berkeley advanced reconstruction toolbox (BART)⁴⁵ was used with patch size $p=8$ and a maximum number of optimization iterations of 80. The optimal value of regularization parameter $\lambda_{\text{LLR}}=2.06$ was chosen via grid search to minimize the reconstructed flow field residual $\|\hat{\mathbf{v}} - \mathbf{v}^*\|_2$ averaged over the manually segmented aorta on the retrospectively $12 \times$ undersampled acquisition. Since the original VN¹⁵ was proposed for magnitude reconstruction of 2D and 3D data, we introduced the following modifications for the presented 4D flow evaluation: (i) 3D filters were grouped into 4 banks as in FlowVN (see Supplemental Algorithm 1), (ii) ℓ_1 -norm of reconstruction was optimized (i.e. equation (3) with $\tau \rightarrow +\infty$). We refer to this architecture as ‘‘HamVN’’, the number of network layers, filters and control knots were the same as in FlowVN.

Retrospective Study

For simulated retrospective undersampling experiments, we used $2 \times$ PI data and simulate a pseudo-radial Golden angle sampling pattern³⁹ with acceleration factors of 6 to 22.

For each undersampling factor we evaluated the nRMSE of the image magnitude, the relative error (RelErr) of velocity magnitudes inside the aorta and the angular error (AngErr) of the

estimated velocity vectors:

$$\begin{aligned} \text{nRMSE}(\mathbf{a}, \mathbf{a}^*) &= \sqrt{\sum_i^N \frac{(a_i - a_i^*)^2}{N \max_j (a_j^*)^2}}, \\ \text{RelErr}(\mathbf{a}, \mathbf{a}^*) &= \frac{\|\mathbf{a} - \mathbf{a}^*\|_2}{\|\mathbf{a}^*\|_2}, \\ \text{AngErr}(\mathbf{u}, \mathbf{v}) &= \arccos\left(\frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|_2 \|\mathbf{v}\|_2}\right). \end{aligned} \tag{11}$$

Additionally, we report the structural similarity index (SSIM)⁴⁶ with $\sigma_{\text{SSIM}}=1.5$ on the reconstructed magnitude images.

Prospective Study

Using manual aorta segmentations we compute flow over cross sections of the aorta by integrating velocity components projected onto the cross section normal. The peak flow is then defined as the maximal flow over cardiac phases for a given cross section. Moreover, we calculate peak through-plane velocity defined as maximum velocity projection across cross sections of the aorta over cardiac phases.

To quantify agreement with the reference $2\times$ PI reconstruction, we performed Bland-Altman analysis⁴⁷ of peak flow and peak through-plane velocities.

Data and code availability

The code for the network training and inference used in this study and network weights are available on CodeOcean together with volunteer data: <https://codeocean.com/capsule/0115983/tree>⁴⁸. The code for analysis is available on CodeOcean: <https://codeocean.com/capsule/2587940/tree>⁴⁹.

References

- [1] Markl, M., Frydrychowicz, A., Kozerke, S., Hope, M. & Wieben, O. 4D flow MRI. *J. Mag. Res. Imag.* **36**, 1015–1036 (2012).
- [2] Feinberg, D., Hale, J., Watts, J., Kaufman, L. & Mark, A. Halving MR imaging time by conjugation: demonstration at 3.5 kG. *Radiol.* **161**, 527–531 (1986).
- [3] Szarf, G. *et al.* Zero filled partial fourier phase contrast MR imaging: in vitro and in vivo assessment. *J. Mag. Res. Imag.* **23**, 42–49 (2006).
- [4] Walheim, J., Gotschy, A. & Kozerke, S. On the limitations of partial fourier acquisition in phase-contrast MRI of turbulent kinetic energy. *Mag. Res. in Med.* **81**, 514–523 (2019).
- [5] Pruessmann, K., Weiger, M., Scheidegger, M. & Boesiger, P. SENSE: sensitivity encoding for fast MRI. *Mag. Res. in Med.* **42**, 952–962 (1999).
- [6] Wiesinger, F., Boesiger, P. & Pruessmann, K. P. Electrodynamics and ultimate SNR in parallel MR imaging. *Mag. Res. in Med.* **52**, 376–390 (2004).
- [7] Lustig, M., Donoho, D. & Pauly, J. M. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Mag. Res. in Med.* **58**, 1182–1195 (2007).
- [8] Kim, D. *et al.* Accelerated phase-contrast cine MRI using k-t SPARSE-SENSE. *Mag. Res. in Med.* **67**, 1054–1064 (2012).
- [9] Valvano, G. *et al.* Accelerating 4D flow MRI by exploiting low-rank matrix structure and hadamard sparsity. *Mag. Res. in Med.* **78**, 1330–1341 (2017).
- [10] Bollache, E. *et al.* k-t accelerated aortic 4D flow MRI in under two minutes: feasibility and impact of resolution, k-space sampling patterns, and respiratory navigator gating on hemodynamic measurements. *Mag. Res. in Med.* **79**, 195–207 (2018).

- [11] Walheim, J., Dillinger, H. & Kozerke, S. Multipoint 5D Flow Cardiovascular Magnetic Resonance - Accelerated Cardiac- and Respiratory-Motion Resolved Mapping of Mean and Turbulent Velocities. *J. Cardiovasc. Magn. Res.* (2019).
- [12] Ma, L. E. *et al.* Aortic 4D flow MRI in 2 minutes using compressed sensing, respiratory controlled adaptive k-space reordering, and inline reconstruction. *Mag. Res. in Med.* (2019).
- [13] Rich, A. *et al.* A bayesian approach for 4D flow imaging of aortic valve in a single breath-hold. *Mag. Res. in Med.* **81**, 811–824 (2019).
- [14] Zhang, T., Pauly, J. M. & Levesque, I. R. Accelerating parameter mapping with a locally low rank constraint. *Mag. Res. in Med.* **73**, 655–661 (2015).
- [15] Hammernik, K. *et al.* Learning a variational network for reconstruction of accelerated MRI data. *Mag. Res. in Med.* **79**, 3055–3071 (2018).
- [16] Mardani, M. *et al.* Deep generative adversarial neural networks for compressive sensing MRI. *IEEE Trans. Med. Imag.* **38**, 167–179 (2018).
- [17] Zhu, B., Liu, J. Z., Cauley, S. F., Rosen, B. R. & Rosen, M. S. Image reconstruction by domain-transform manifold learning. *Nature* **555**, 487 (2018).
- [18] 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 Trans. Med. Imag.* **37**, 491–503 (2017).
- [19] Maier, A. K. *et al.* Learning with known operators reduces maximum error bounds. *Nat. Mach. Intel.* **1**, 373–380 (2019).
- [20] Beck, A. & Teboulle, M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imag. Sc.* **2**, 183–202 (2009).
- [21] Antun, V., Renna, F., Poon, C., Adcock, B. & Hansen, A. C. On instabilities of deep learning in image reconstruction—does AI come at a cost? *arXiv preprint arXiv:1902.05300* (2019).
- [22] Yang, G. *et al.* DAGAN: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. *IEEE Trans. Med. Imag.* **37**, 1310–1321 (2017).
- [23] Quan, T., Nguyen-Duc, T. & Jeong, W. Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss. *IEEE Trans. Med. Imag.* **37**, 1488–1497 (2018).
- [24] Narnhofer, D., Hammernik, K., Knoll, F. & Pock, T. Inverse GANs for accelerated MRI reconstruction. In *Wav. and Spars.*, vol. 11138, 111381A (2019).
- [25] Zhang, S., Block, K. & Frahm, J. Magnetic resonance imaging in real time: advances using radial FLASH. *J. Mag. Res. Imag.* **31**, 101–109 (2010).
- [26] Landweber, L. An iteration formula for Fredholm integral equations of the first kind. *Am. J. of Math.* **73**, 615–624 (1951).
- [27] Liu, Y. & Lew, M. S. Learning relaxed deep supervision for better edge detection. In *CVPR*, 231–240 (2016).
- [28] Ravishankar, S. & Bresler, Y. MR image reconstruction from highly undersampled k-space data by dictionary learning. *IEEE Trans. Med. Imag.* **30**, 1028–1041 (2010).
- [29] Caballero, J., Price, A. N., Rueckert, D. & Hajnal, J. V. Dictionary learning and time sparsity for dynamic MR data reconstruction. *IEEE Trans. Med. Imag.* **33**, 979–994 (2014).
- [30] Lee, D., Yoo, J. & Ye, J. C. Deep artifact learning for compressed sensing and parallel MRI. *arXiv preprint arXiv:1703.01120* (2017).
- [31] LeCun, Y., Touresky, D., Hinton, G. & Sejnowski, T. A theoretical framework for back-propagation. In *Proc. Connect. Mod.*, vol. 1, 21–28 (1988).

- [32] Abadi, M. *et al.* Tensorflow: A system for large-scale machine learning. In *Proc. USENIX*, 265–283 (2016).
- [33] Domke, J. Generic methods for optimization-based modeling. In *Art. Int. and Stat.*, 318–326 (2012).
- [34] Sun, J., Li, H., Xu, Z. *et al.* Deep ADMM-Net for compressive sensing MRI. In *NIPS*, 10–18 (2016).
- [35] Jin, K. H., McCann, M. T., Froustey, E. & Unser, M. Deep convolutional neural network for inverse problems in imaging. *IEEE Trans. on Imag. Proc.* **26**, 4509–4522 (2017).
- [36] Vishnevskiy, V., Sanabria, S. J. & Goksel, O. Image reconstruction via variational network for real-time hand-held sound-speed imaging. In *Intl. W. on Mach. Learn. for Med. Imag. Recon.*, 120–128 (2018).
- [37] Vishnevskiy, V., Rau, R. & Goksel, O. Deep variational networks with exponential weighting for learning computed tomography. In *MICCAI*, 310–318 (2019).
- [38] Cuppen, J. & van Est, A. Reducing MR imaging time by one-sided reconstruction. *Mag. Res. Imag.* **5**, 526–527 (1987).
- [39] Winkelmann, S., Schaeffter, T., Koehler, T., Eggers, H. & Doessel, O. An optimal radial profile order based on the Golden Ratio for time-resolved MRI. *IEEE Trans. Med. Imag.* **26**, 68–76 (2006).
- [40] Zhang, T., Pauly, J. M., Vasanawala, S. S. & Lustig, M. Coil compression for accelerated imaging with cartesian sampling. *Mag. Res. in Med.* **69**, 571–582 (2013).
- [41] Uecker, M. *et al.* ESPIRiT — an eigenvalue approach to autocalibrating parallel MRI: where SENSE meets GRAPPA. *Mag. Res. in Med.* **71**, 990–1001 (2014).
- [42] Bernstein, M. A. *et al.* Concomitant gradient terms in phase contrast MR: analysis and correction. *Mag. Res. in Med.* **39**, 300–308 (1998).
- [43] Busch, J., Giese, D. & Kozerke, S. Image-based background phase error correction in 4D flow MRI revisited. *J. of MRI* **46**, 1516–1525 (2017).
- [44] Walker, P. G. *et al.* Semiautomated method for noise reduction and background phase error correction in MR phase velocity data. *J. Mag. Res. Imag.* **3**, 521–530 (1993).
- [45] Tamir, J. I., Ong, F., Cheng, J. Y., Uecker, M. & Lustig, M. Generalized magnetic resonance image reconstruction using the Berkeley advanced reconstruction toolbox. In *ISMRM W. on Dat. Sampl. and Imag. Recon.* (2016).
- [46] Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. on Imag. Proc.* **13**, 600–612 (2004).
- [47] Altman, D. G. & Bland, J. M. Measurement in medicine: the analysis of method comparison studies. *J. R. Stat. Soc.* **32**, 307–317 (1983).
- [48] Vishnevskiy, V., Walheim, J. & Kozerke, S. FlowVN: deep variational network for rapid 4D flow MRI reconstruction. *CodeOcean* URL <https://codeocean.com/capsule/0115983/tree>. (2020).
- [49] Vishnevskiy, V., Walheim, J. & Kozerke, S. FlowVN: analysis. *CodeOcean* URL [CodeOcean\url{https://codeocean.com/capsule/5994453/tree}](https://codeocean.com/capsule/5994453/tree). (2020).

Acknowledgements

The authors acknowledge funding from the European Unions Horizon 2020 research and innovation program under grant agreement No 668039 and under EuroStars UNIFORM as well as funding of the Platform for Advanced Scientific Computing of the Council of the Federal Institutes of Technology (ETH Board), Switzerland.

Author contributions

J.W., V.V. and S.K. conceived the study. V.V. implemented machine learning reconstruction algorithms. J.W. conducted MR acquisition experiments and data preprocessing. J.W. and V.V. analyzed experimental data under supervision of S.K. All authors discussed the results and contributed to writing the manuscript.

Competing interests

The authors declare no competing interests.

1 Supplementary Material

Algorithm 1 Proposed variational reconstruction network model FlowVN.

Input: $\mathbf{B} \in \mathbb{C}^{N_s \times N_t}$ — zero-filled k-space samples, $\mathbf{M} \in \{0, 1\}^{N_s \times N_t}$ — undersampling mask

Parameters: $\Theta = \{\phi_{ud}^{(k)}, \phi_{ur}^{(k)}, \phi_d^{(k)}, \phi_{r,in}^{(k)}, \mathbf{D}_{in}^{(k)}, \alpha^{(k)}, \alpha^{(0)}\}_{i=1, \dots, N_f, k=1, \dots, K, n=1, \dots, 4}$

$\overline{\mathbf{M}} \leftarrow \frac{1}{N_s N_t} \sum_{i \leq N_s, j \leq N_t} M_{ij}$
 $\mathbf{P}^{(0)} \leftarrow \alpha^{(0)} \mathbf{E}^H \mathbf{B}, \quad \mathbf{S}^{(0)} \leftarrow \mathbf{0}$

for $k := 0$ to $K - 1$

$$\mathbf{G}^{(k)} \leftarrow \overbrace{\varphi_{ud}^{(k)} \{\overline{\mathbf{M}}\} \mathbf{E}^H \left(\mathbf{M} \odot \varphi_d^{(k)} \left\{ \mathbf{M} \odot \left(\mathbf{E} \mathbf{P}^{(k)} - \mathbf{B} \right) \right\} \right)}^{\text{data consistency term}} + \overbrace{\varphi_{ur}^{(k)} \{\overline{\mathbf{M}}\} \sum_{i \leq N_f, n \leq 4} \left(\mathbf{D}_{in}^{(k)} \right)^T \varphi_{r,in}^{(k)} \left\{ \mathbf{D}_{in}^{(k)} \text{vec} \left(\mathbf{P}^{(k)} \right) \right\}}^{\text{regularization term}}$$

$$\mathbf{S}^{(k+1)} \leftarrow \alpha^{(k+1)} \mathbf{S}^{(k)} + \mathbf{G}^{(k)}$$

$$\mathbf{P}^{(k+1)} \leftarrow \mathbf{P}^{(k)} - \mathbf{S}^{(k+1)}$$

Output: reconstructed image $\mathcal{V}(\mathbf{B}, \mathbf{M}; \Theta) \stackrel{\text{def}}{=} \mathbf{P}^{(K)}$

Table 2: Comparison of reconstruction errors (\pm standard deviation) of 7 retrospectively under-sampled acquisitions of healthy subjects for different acceleration factors R . For each row the best performing method is highlighted in bold. P-values are obtained by comparing error metrics distributions of LLR and corresponding VN architectures using the two-sided Wilcoxon signed-rank test.

	Image Magnitude nRMSE [%]			Velocity Magnitude RelErr [%]			Mean Flow AngErr [deg]		
	LLR	HamVN	FlowVN	LLR	HamVN	FlowVN	LLR	HamVN	FlowVN
$R=6$	1.9 \pm 0.3	2.3 \pm 0.2	1.7 \pm 0.2	9.4 \pm 1.9	12.8 \pm 2.5	8.7 \pm 2.4	7.0 \pm 1.8	9.6 \pm 2.6	6.9 \pm 1.9
p-val.		<0.0001	0.0017		<0.0001	0.0225		<0.0001	0.1956
$R=8$	2.0 \pm 0.3	2.5 \pm 0.2	1.9 \pm 0.2	10.5 \pm 2.3	13.5 \pm 2.8	10.0 \pm 2.8	7.9 \pm 2.0	10.5 \pm 2.9	8.0 \pm 2.3
p-val.		<0.0001	0.0076		<0.0001	0.0790		<0.0001	0.2228
$R=10$	2.2 \pm 0.3	2.8 \pm 0.3	2.2 \pm 0.2	11.8 \pm 2.2	14.9 \pm 3.0	11.4 \pm 2.8	8.9 \pm 2.2	11.5 \pm 3.1	9.2 \pm 2.5
p-val.		<0.0001	0.1418		<0.0001	0.0385		<0.0001	0.0204
$R=12$	2.5 \pm 0.3	3.0 \pm 0.3	2.4 \pm 0.2	13.2 \pm 2.4	16.4 \pm 2.9	12.4 \pm 2.4	10.2 \pm 2.7	12.7 \pm 3.4	10.3 \pm 2.8
p-val.		<0.0001	0.0256		<0.0001	0.0003		<0.0001	0.0225
$R=14$	2.8 \pm 0.3	3.3 \pm 0.3	2.6 \pm 0.3	15.1 \pm 2.5	18.0 \pm 3.5	14.0 \pm 2.9	11.3 \pm 2.9	13.6 \pm 3.8	11.4 \pm 3.2
p-val.		<0.0001	<0.0001		<0.0001	<0.0001		<0.0001	0.1166
$R=16$	3.3 \pm 0.4	3.6 \pm 0.3	2.9 \pm 0.3	17.4 \pm 2.3	20.2 \pm 3.2	15.5 \pm 3.3	12.2 \pm 3.1	14.3 \pm 4.0	12.3 \pm 3.4
p-val.		<0.0001	<0.0001		<0.0001	<0.0001		<0.0001	0.7785
$R=18$	3.6 \pm 0.4	3.8 \pm 0.4	3.0 \pm 0.3	19.0 \pm 2.7	22.4 \pm 2.9	16.9 \pm 2.9	13.3 \pm 3.4	15.3 \pm 4.2	13.0 \pm 3.6
p-val.		<0.0001	<0.0001		<0.0001	<0.0001		<0.0001	0.4799
$R=20$	3.8 \pm 0.4	4.0 \pm 0.4	3.2 \pm 0.3	19.9 \pm 3.4	23.1 \pm 4.1	17.5 \pm 3.4	13.7 \pm 3.7	15.5 \pm 4.4	13.6 \pm 3.9
p-val.		0.0211	<0.0001		<0.0001	<0.0001		<0.0001	0.0073
$R=22$	4.2 \pm 0.5	4.1 \pm 0.4	3.4 \pm 0.3	21.9 \pm 2.8	25.3 \pm 3.6	19.2 \pm 3.3	14.8 \pm 3.6	16.1 \pm 4.5	14.3 \pm 3.8
p-val.		0.9551	<0.0001		<0.0001	<0.0001		<0.0001	0.0005

Table 3: Quantitative prospective comparison of VN architectures and modifications proposed in this paper using data from 7 healthy volunteers ($12.4 \leq R \leq 13.8$). Peak velocities and peak flow are assessed over manually segmented aorta slices and reported relative to average values of the corresponding reference scan (\pm standard deviation). P-values are obtained by comparing error metrics distributions of reference PI data and the corresponding reconstruction method using the two-sided Wilcoxon signed-rank test. HamVN-3D performs 3D filtering of yzt dimensional data (assuming that data is fully sampled in the x dimension). “HamVN-3D, single coil” was trained to reconstruct each coil image separately and then to be combined using provided sensitivity maps. Modifications are accumulated as presented and, eventually, yield the FlowVN configuration.

Method	Peak Velocity Error[%]	p-value	Peak Flow Error[%]	p-value
HamVN-3D, single coil	19.38 \pm 19.01	<0.0001	21.37 \pm 15.11	<0.0001
HamVN-3D	6.44 \pm 13.64	<0.0001	3.73 \pm 12.12	0.0020
HamVN	3.64 \pm 10.10	<0.0001	1.90 \pm 9.69	0.3345
+linear activation functions	3.18 \pm 10.35	0.0002	1.65 \pm 10.70	0.3833
+unroll GD with momentum	2.95 \pm 9.89	0.0004	1.47 \pm 9.82	0.4749
+data activation functions	2.29 \pm 10.12	0.0213	0.63 \pm 9.74	0.7957
+US-depending modulation	1.98 \pm 8.43	0.0645	0.47 \pm 10.30	0.9272
+exp. weighting (FlowVN)	1.59 \pm 9.65	0.0875	0.05 \pm 9.79	0.9550

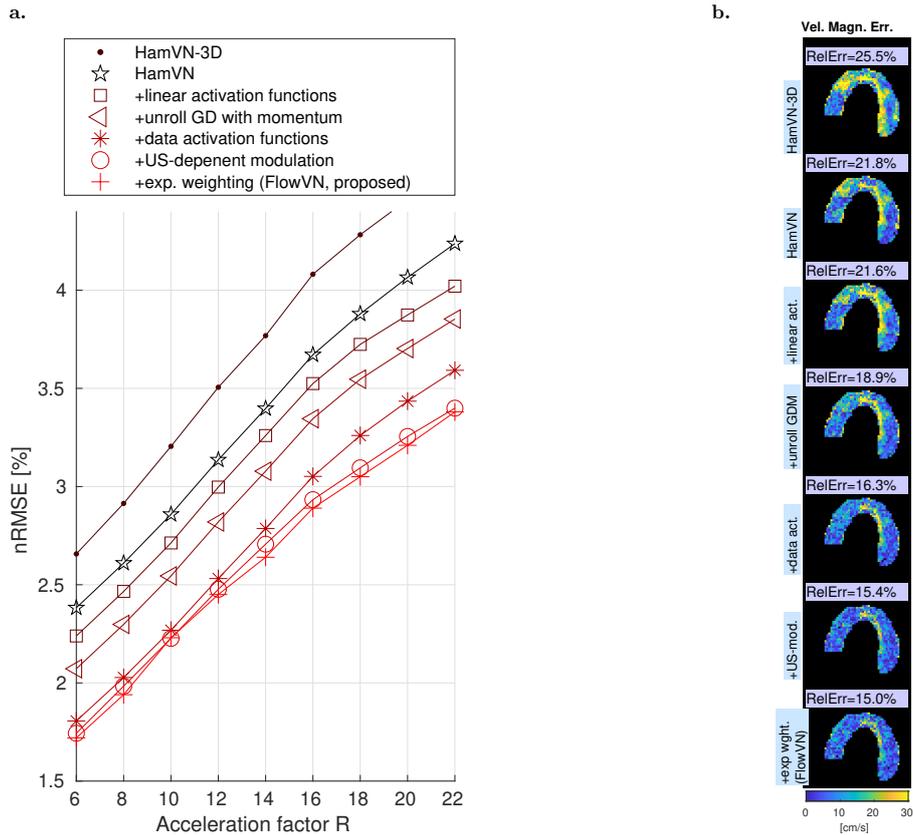


Figure 6: **Retrospective ablation study.** **a**, Comparison of VN architectures and modifications proposed in this paper, using retrospectively undersampled data from 7 healthy volunteers for different acceleration factors R . HamVN-3D performs 3D filtering of yzt dimensional data (assuming that data is fully sampled in the x dimension). Modifications are accumulated as presented in the legend and, eventually, yield the FlowVN configuration. Undersampling-dependent modulation improves reconstruction quality only for higher acceleration factors. Note that the layer-wise exponential weighting of the reconstruction loss does not increase model complexity and, therefore, does not improve reconstruction accuracy for retrospective acceleration simulations. **b**, Velocity magnitude error maps for $R=16$ shown at systolic peak flow.