Separating Aliasing from Flow: Temporal-Frequency Regularization for Unrolled 4D flow MRI Reconstruction

Ji-Hoon Jung[1,2][0009-0002-7416-6727]

Jihun Kang, PhD[1][0000-0001-5455-6651]

Hojin Ha, PhD[3][0000-0001-6476-2889]

June-Goo Lee, PhD[1,2][0000-0002-1380-6682]

[1] Biomedical Engineering Research Center, Asan Institute for Life Sciences, Asan Medical Center, Seoul, South Korea
[2] Department of Biomedical Engineering, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea
[3] Department of Smart Health Science and Technology, Kangwon National University, Chuncheon, South Korea
junegoo.lee@gmail.com (corresponding author)

Abstract. Four-dimensional (4D) flow MRI encodes three velocity components over the cardiac cycle in a 10–20 minute scan; reaching the recommended 5–10 minutes calls for high undersampling, where velocity-derived measures degrade far faster than image magnitude. We present XF-FlowNet, our entry to the CMRx4DFlow2026 challenge, which keeps the widely used serial unrolled cascade skeleton — a residual regularizer followed by a data-consistency (DC) step, repeated K times — and adds three components targeted at the velocity field: refinement in the temporal-frequency (x–f) domain with an explicit $|f|$ positional encoding, conditioning of the regularizer on the sampling mask spectrum and the current DC residual, and a zero-initialized gated conjugate-gradient DC block. On 30 cases at five acceleration rates ($R = 10$–50, 150 reconstructions) it improves over the same skeleton without these components by $+0.0156$ SSIM, -0.0113 nRMSE, -0.0611 RelErr and $-5.20°$ angular error, with every metric improving in essentially all 150 reconstructions ($p < 0.001$, paired Wilcoxon). Run with half the cascades, the same weights clear the challenge quality gate on all 32 official cases while reducing inference time by 40 %.

Keywords: 4D flow MRI · Velocity encoding · Image reconstruction · Temporal Fourier transform · Data consistency · Deep learning.

1 Introduction

4D flow MRI measures a three-directional, time-resolved velocity field over an entire 3D volume, from which flow velocity, wall shear stress and vorticity follow

across the cardiac cycle [3, 6, 11], supporting the assessment of aortic aneurysm, stenosis and dissection [3, 5]. Acquisition time is what limits routine use: current acceleration reaches 10–20 minutes at low to medium resolution [5], whereas the 2023 consensus statement recommends staying within 5–10 minutes [3]. Because the acquisition spans three spatial dimensions, time and four velocity encodings, the burden is large but so is the space available for undersampling [5].

What makes this regime hard is that the two families of quality measures come apart. The quantity of interest is the phase difference between velocity encodings, not the image magnitude [12], and acceleration biases it — k–t accelerated 4D flow MRI underestimated peak flow more strongly than parallel imaging in a head-to-head validation [4]. Velocity-derived metrics are therefore the discriminative criterion here, and the method must be driven by the velocity field rather than by image appearance.

Temporal-frequency (x–f) sparsity is the classical answer to that problem and has long been used to accelerate flow imaging [2, 7], and x–f regularization inside a deep unrolled network has been demonstrated for cine [15, 16]. Deep-learning reconstruction of 4D flow MRI, however, has so far refined in the image domain [9, 14, 20]. To our knowledge this work is the first to place the regularizer of an unrolled 4D flow MRI reconstruction in the x–f domain.

We develop and evaluate the method within the CMRx4DFlow2026 challenge, whose two regular tasks separate the two demands stated above [5]. Task 1 assesses robustness across clinical centres and scanners at 10×–50× acceleration and ranks by the sum of ranks over four metrics: SSIM and nRMSE on the magnitude image, relative velocity error (RelErr) and angular error (AngErr) on the velocity field. Task 2 admits a submission only if its complex-difference error is below the challenge baseline [20] and then ranks by reconstruction time on standardized hardware. This work targets Task 1 and uses Task 2 to characterize the accuracy–speed behaviour of the same trained model.

Our contributions are as follows.

1. **A conditioned x–f regularizer for 4D flow MRI.** We place the regularizer of an unrolled 4D flow MRI reconstruction in the temporal-frequency domain — to our knowledge the first such regularizer for 4D flow MRI — and demonstrate its effect in a controlled ablation.
2. **Half the compute within the quality gate.** The same trained weights run with five instead of ten cascades cut reconstruction from 22.50 ± 5.37 to 13.46 ± 3.83 s per case ($-40\,\%$) and still beat the challenge baseline [20] on the official set, cutting its complex-difference error by $41.9\,\%$.
3. **Leading accuracy on the challenge leaderboard.** Run with all ten cascades, the method ranks first on the Task 1 validation leaderboard.

2 Method

Fig. 1 shows the complete XF-FlowNet pipeline. Panel (a) is the outer loop, panel (b) one cascade, and panel (c) the regularizer that this work contributes.

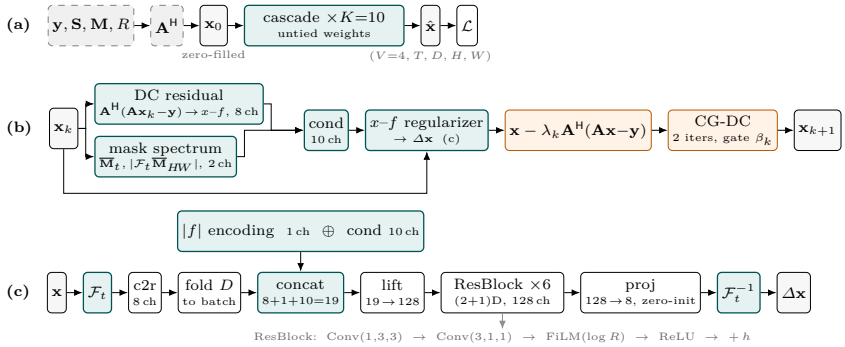

Fig. 1. Architecture. **(a)** $\mathbf{x}_0 = \mathbf{A}^{\mathsf{H}}\mathbf{y}$ enters K=10 untied cascades and leaves as $\hat{\mathbf{x}}$. **(b)** One cascade builds the 10-channel conditioning tensor, refines with the x–f regularizer, then applies a learned-step-size gradient step and a gated conjugate-gradient step. **(c)** The regularizer transforms to x–f, concatenates the $|f|$ encoding and conditioning (8+1+10=19), refines with six (2+1)D blocks at 128 channels, and projects back through a zero-initialized convolution.

2.1 Problem Formulation

Let $\mathbf{x} \in \mathbb{C}^{V\times T\times D\times H\times W}$ be the complex image series with $V = 4$ velocity encodings ($v = 0$ is the reference), T cardiac phases, readout axis D and phase-encoding axes (H, W), and let $\mathbf{y} = \mathbf{A}\mathbf{x} + \boldsymbol{\eta}$ be the measurements. The forward operator $\mathbf{A} = \mathbf{MFS}$ applies right to left — coil sensitivities, Fourier transform along (H, W), then the sampling mask — together with the sensitivity maps, the aortic segmentation masks and the zero-filled SENSE initialization $\mathbf{x}_0 = \mathbf{A}^{\mathsf{H}}\mathbf{y}$ (Fig. 1a), is taken unchanged from the challenge pipeline [5]; none of it is learned.

2.2 Unrolled Skeleton

We use the gradient-descent data-consistency sensitivity network ($\mathrm{SN}_{\mathrm{GD}}$) of Hammernik et al. [8] as the skeleton: each of the K=10 cascades applies a residual regularizer and then a data-consistency step to *its output*,

$$\mathbf{z}_k = \mathbf{x}_k + \mathcal{D}_{\theta_k}(\mathbf{x}_k), \tag{1}$$

$$\mathbf{x}_{k+1} = \mathbf{z}_k - \lambda_k\, \mathbf{A}^{\mathsf{H}}(\mathbf{A}\mathbf{z}_k - \mathbf{y}), \tag{2}$$

with a per-cascade learned step size $\lambda_k = \mathrm{softplus}(\omega_k)$ and untied weights. Because $\mathbf{A}^{\mathsf{H}}\mathbf{A}$ is not idempotent in the multi-coil case, (2) is a gradient step rather than a projection, so the correction spreads beyond the sampled locations. The acceleration rate R is supplied with the data and is known at inference; every regularizer is conditioned on it. Everything below replaces $\mathcal{D}_{\theta_k}$ and appends one further data-consistency block; the skeleton itself is unchanged.

2.3 Temporal-Frequency (x–f) Regularizer

Aliasing from k–t undersampling and cardiac dynamics are entangled in the image domain: along t both appear as frame-to-frame variation, so a temporal

convolution suppresses aliasing only by also suppressing the flow waveform. A Fourier transform along t separates them — periodic flow concentrates in a few temporal frequencies, aliasing spreads as a broad floor — which is what classical k–t compressed sensing exploits [10]. We therefore let the regularizer act on $X = \mathcal{F}_t\mathbf{x}$ and transform the result back (Fig. 1c):

$$\mathcal{D}_{\theta_k}(\mathbf{x}) = \mathcal{F}_t^{-1}\, g_{\theta_k}\big(\mathcal{F}_t\mathbf{x},\ \mathbf{e},\ \mathbf{c}_k,\ R\big). \tag{3}$$

Real-valued packing. The complex x–f volume is split into real and imaginary parts, giving $2V = 8$ channels. Folding the readout axis D into the batch dimension is exact, because $\mathbf{F}$ acts only on (H, W) and the coil combination is pointwise in D; a five-dimensional volume is thus handled by 3D convolutions over (f, H, W) at a fraction of the memory.

$|f|$ *positional encoding* $\mathbf{e}$. A convolution is shift-invariant along f, so the network cannot tell the temporal DC component from high frequencies unless told where it is — yet that is exactly the distinction the x–f representation exists to exploit. We append one channel carrying the normalized $|f| \in [0, 1]$ coordinate; this single change reduced the temporal averaging seen with an image-domain regularizer.

Backbone. Six residual blocks at 128 channels operate at full resolution, with no down- or upsampling, so no temporal frequency is discarded by pooling. Each block factorizes the 3D convolution [19] into a spatial $(1, 3, 3)$ and a temporal-frequency $(3, 1, 1)$ convolution. The factorization follows from the change of domain: a joint spatio-temporal kernel is natural where motion couples the axes, but the Fourier transform along t weakens that coupling, leaving a per-voxel spectral profile along f and spatial structure within each frequency. Separating the two also costs 12 rather than 27 multiplies per channel pair, which is what makes 128 channels at full resolution affordable. The spatial dilations are $(1, 2, 4, 2, 1, 1)$ across the six blocks, which gives a 27-pixel spatial receptive field without any downsampling; the f convolution is never dilated. Each pair is followed by feature-wise linear modulation [13] whose scale and shift come from an MLP on $\log R$, a ReLU, and a residual addition. Unlike R(2+1)D [19] we place no nonlinearity between the two convolutions, so each pair stays a separable linear kernel — the form the decoupled x–f representation calls for. The $128 \to 8$ output convolution is zero-initialized, so at initialization $\mathcal{D}_{\theta_k} \equiv 0$ and the network reduces exactly to K-step gradient descent on the data term.

2.4 Conditioning

The regularizer is given ten further input channels (Fig. 1b), both carried on the same (f, H, W) grid as the features they modulate.

Data-consistency residual (8 channels). $\mathcal{F}_t\mathbf{A}^{\mathsf{H}}(\mathbf{A}\mathbf{x}_k - \mathbf{y})$, packed as $2V$ real channels, marks per voxel and per temporal frequency where the current estimate disagrees with the measurements, i.e. which part of the estimate is extrapolated rather than measured.

Mask spectrum (2 channels). The mask is a k-space object, so rather than feed it directly we supply two reduced summaries: the time-averaged sampling density $\overline{\mathbf{M}}_t$, which keeps the k-space (H, W) grid and is constant along f, and $|\mathcal{F}_t \overline{\mathbf{M}}_{H,W}|$, the spatially averaged temporal spectrum, which is the aliasing kernel along the f axis on which the backbone convolves. Both are normalized to $[0, 1]$ and broadcast over the folded readout axis.

2.5 Gated Conjugate-Gradient Data Consistency

At $R = 50$ only about 2 % of each frame is measured, and the single gradient step of (2) propagates those few samples too weakly. We append a proximal data-consistency block that solves

$$(\mathbf{A}^{\mathsf{H}}\mathbf{A} + \mu_k \mathbf{I})\,\mathbf{x} = \mathbf{A}^{\mathsf{H}}\mathbf{y} + \mu_k \mathbf{x}_{k+1} \tag{4}$$

with two conjugate-gradient iterations, as in MoDL [1], which lets the measured samples propagate globally through the coil coupling. The block enters through a learned gate, $\mathbf{x} \leftarrow \mathbf{x} + \beta_k(\mathbf{x}_{\mathrm{CG}} - \mathbf{x})$, with β_k zero-initialized so that it is exactly the identity at initialization and is switched on only where it helps. After training, β_k settles near 1.2 in all ten cascades.

2.6 Training Objective

The loss mirrors the four challenge metrics [5] (SSIM, nRMSE, squared relative velocity error and an angular term) plus a small ℓ_1 term. Each term is divided by a running exponential-moving-average of its own scale so that the four contribute comparably; without this normalization the angular term dominates and SSIM stagnates. Velocity terms are restricted to the segmentation mask; magnitude terms use a five-voxel dilation of it, weighted 0.1 outside the dilation.

The score averages the five acceleration rates equally although they are not equally hard, so each sample's loss is additionally scaled by a weight that rises with the running loss at its rate — a soft group-DRO [17] over R, read before the current sample enters the average so that it cannot feed back on itself.

2.7 Training Configuration

One optimizer step consumes a single sample: a slab of $D = 5$ adjacent readout slices with all $V = 4$ encodings and $T = 15$ cardiac phases, whose start is drawn at random and wrapped cyclically to respect the periodicity of the cycle. The batch size is one because the phase-encoding matrix (H, W) differs between cases and scanners, so samples cannot be stacked; enumerating every slice position gives 12,246 samples per epoch. Each sample draws its rate uniformly from $R \in \{10, 20, \ldots, 50\}$ with a freshly generated k–t Gaussian mask, so no pattern is seen twice; validation uses full volumes and, for each (case, R) pair, one mask drawn from the same generator under a fixed seed, so every configuration is scored on identical patterns. Optimization uses Adam at 10^{-4}, a cosine schedule and gradient clipping at 1.0.

3 Experiments and Results

Data, implementation and ablation protocol. We hold out 30 of the 138 CMRx-4DFlow2026 training cases as an internal validation set, stratified over six centre/scanner combinations, and evaluate each at all five rates $R \in \{10, 20, \ldots, 50\}$, giving 150 reconstructions. Scoring uses the official metric implementation, with MSAC background-phase correction on the ground truth, verified against the organizers' script. The model uses K=10 cascades, 128 channels and six residual blocks (12.6M parameters); all training runs on a single NVIDIA H200 NVL and all inference on an NVIDIA RTX A6000, the challenge reference hardware, where a full volume takes 22.50 ± 5.37 s. Ablation configurations are trained from scratch under a common budget of 150,000 steps — the submitted model is trained for 600,000 — so that the comparison reflects the component under test and not a difference in training length, and a single shared data loader delivers *identical* mini-batches and masks to every configuration at every step, making all comparisons paired at the sample level.

3.1 Component Ablation

Table 1 reports mean ± standard deviation over the $n = 150$ reconstructions; the deviation is large because it aggregates case difficulty and five very different rates, so significance uses a paired Wilcoxon signed-rank test on the same 150 pairs.

These 150 reconstructions come from 30 subjects at five rates each, so they are not 150 independent observations. We therefore repeated every test at the subject level, averaging the five rates of each subject into 30 paired values per metric: every comparison against B0 remains significant (all $p < 0.001$; largest $p = 9.5 \times 10^{-4}$, B1 angular error), and for B3 and B4 all 30 subjects improve on all four metrics (two-sided exact $p = 1.9 \times 10^{-9}$). The conclusions of Table 1 are unchanged under this subject-level analysis.

The conditioning is carried on the same (f, H, W) grid the backbone convolves over, and the mask spectrum has no image-domain counterpart (Sec. 2.4); the two are one design, and B1 measures the representation stripped of it.

Since the conditioning is defined inside the x–f domain, B2 — not B1 — is the x–f regularizer as proposed: it improves all four metrics over B0, with angular error reduced by 3.30°. B1, which strips the conditioning away, is the lower bound of the representation rather than the proposed design. The gated CG block (B3) improves all four metrics as well, with angular error reduced by 4.48°. The two branches are complementary: neither alone reaches XF-FlowNet (B4), which is best on all four metrics, and each adds about +0.005 SSIM on top of the other.

3.2 Task 1 — Reconstruction Accuracy

We submit XF-FlowNet with all ten cascades; Table 2 reports its internal and official scores. Running the same weights with only the first five cascades —

Table 1. Component ablation on the internal validation set (n = 150: 30 cases × 5 acceleration rates). Mean ± standard deviation. *** denotes $p < 0.001$ against B0 (paired Wilcoxon signed-rank on the same 150 sample pairs; unchanged at the subject level, n=30, Sec. 3.1). All configurations are trained from scratch for 150,000 steps on identical mini-batches.

	x–f	cond	CG-DC	Params	SSIM ↑	nRMSE ↓	RelErr ↓	AngErr (°) ↓
B0				12,505,530	0.9571 ± 0.0310	0.0399 ± 0.0161	0.2813 ± 0.0732	27.33 ± 7.74
B1	✓			12,517,050	0.9584 ± 0.0323***	0.0386 ± 0.0160***	0.2738 ± 0.0757***	26.62 ± 7.83***
B2	✓	✓		12,632,250	0.9680 ± 0.0276***	0.0320 ± 0.0149***	0.2412 ± 0.0712***	24.03 ± 7.17***
B3			✓	12,505,550	0.9679 ± 0.0254***	0.0325 ± 0.0145***	0.2364 ± 0.0715***	22.85 ± 7.22***
B4	✓	✓	✓	12,632,270	**0.9727 ± 0.0252*****	**0.0286 ± 0.0143*****	**0.2202 ± 0.0703*****	**22.13 ± 7.08*****

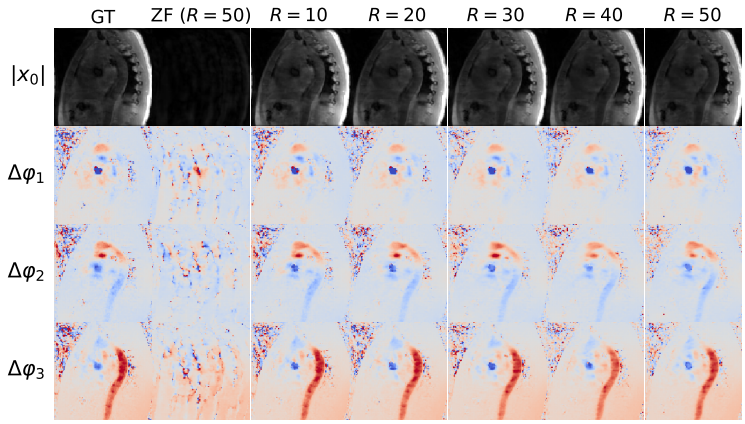

Fig. 2. Worst case of the internal validation set (P004, ranked last at every rate), on the slice where the aortic mask is largest and at the frame with the largest mean phase difference inside it. **Columns:** ground truth, the zero-filled input at R=50, and XF-FlowNet (K=10) at each rate. **Rows:** the magnitude of the reference encoding and the phase differences $\Delta\varphi_k = \angle(x_k x_0^*)$ against it; as in the challenge metric implementation no venc scaling is applied, so these are radians.

the configuration submitted to the efficiency track (Sec. 3.3) — costs little in magnitude quality but a great deal in the velocity field: internally SSIM drops by 1.7 % relative while relative velocity error rises by 71.7 % and angular error by 14.49° (paired, $p < 0.001$), and the per-rate gap in angular error widens from 10.1° at R=10 to 18.3° at R=50, so the removed cascades matter most where the measurements are sparsest. Fig. 2 shows the same magnitude–velocity asymmetry across R: on the hardest case the magnitude barely changes from R=10 to R=50, whereas the phase maps lose the aortic signal. The two tracks therefore separate — Task 1 is ranked by accuracy, Task 2 by inference time once a quality gate is cleared — so the two configurations go to different tracks.

3.3 Task 2 — Efficiency

Because Task 2 ranks by inference time once the quality gate is cleared, we submit the same trained weights executed with five cascades. Table 3 reports the official result: the submission clears the gate on all 32 cases, reduces the

Table 2. Task 1. The K=5 rows share the K=10 weights, executed with the first five cascades. *Top:* internal validation, mean ± standard deviation over $n = 150$ reconstructions and, for K=10, by rate ($n = 30$ each). *Bottom:* official validation leaderboard ($n = 32$, means only); the parenthesized rank is over the 28 accounts' best entries, giving a rank sum of 5 for K=10 — first place. K=5 is our own second entry and is ranked for reference only.

	R	SSIM ↑	nRMSE ↓	RelErr ↓	AngErr (°) ↓
Internal validation ($n = 150$)					
K=10	all	**0.9753 ± 0.0226**	**0.0270 ± 0.0134**	**0.2045 ± 0.0670**	**20.70 ± 6.92**
	10	0.9914 ± 0.0106	0.0142 ± 0.0079	0.1199 ± 0.0334	12.55 ± 3.64
	20	0.9817 ± 0.0168	0.0228 ± 0.0100	0.1776 ± 0.0430	18.12 ± 4.57
	30	0.9740 ± 0.0207	0.0286 ± 0.0113	0.2154 ± 0.0441	21.67 ± 5.01
	40	0.9675 ± 0.0241	0.0329 ± 0.0123	0.2403 ± 0.0439	24.30 ± 5.29
	50	0.9619 ± 0.0257	0.0365 ± 0.0126	0.2691 ± 0.0483	26.86 ± 5.44
K=5	all	0.9590 ± 0.0316	0.0371 ± 0.0165	0.3510 ± 0.1135	35.19 ± 9.84
Official validation leaderboard ($n = 32$)					
K=10	all	0.9754 (1)	0.0275 (2)	0.1990 (1)	18.89 (1)
K=5	all	0.9610 (8)	0.0369 (8)	0.3361 (9)	32.80 (12)

Table 3. Task 2 on the official set ($n = 32$, NVIDIA RTX A6000); both rows share the same weights. "vs. baseline" is the complex-difference error minus the challenge baseline's (0.117568), the gate counts cases beating it, and the speed-up is the mean per-case ratio against K=10 (range 1.40–2.01×). The K=10 row is our own timing of the Task-1 entry on the same hardware, not a Task-2 submission.

	ComplexDiffErr ↓	vs. baseline	Gate	Time/case (s) ↓	Total (min) ↓	Speed-up
XF-FlowNet (K=10)	—	—	—	22.50 ± 5.37	12.00	—
XF-FlowNet (K=5, submitted)	**0.068304**	−0.049264 (−41.9 %)	**32/32**	**13.46 ± 3.83**	**7.18**	**1.70 ± 0.18**×

complex-difference error by 41.9 % relative to the baseline, and removes 40 % of the runtime for a case-wise speed-up of 1.70 ± 0.18×. The speed-up falls short of 2× because a fixed cost — the initial $\mathbf{A}^{\mathsf{H}}\mathbf{y}$ reconstruction, coil combination, the Fourier transforms and tensor I/O — is independent of the number of cascades.

4 Discussion and Conclusion

Two observations follow. First, data consistency remains binding at these rates: the gated CG block improves every one of the 150 reconstructions, fitting the view that at $R = 50$ the difficulty is missing information rather than weak regularization. Second, the temporal-frequency representation and the conditioning that feeds it are one component, not two: stripped of the mask spectrum and the DC residual it yields little (B1), while the conditioned form (B2) matches the CG block on the magnitude metrics — the network appears to need to know which frequencies are measured before a sparse x–f prior can be used safely.

Four limitations bound the scope. Every ablation is a single run with a fixed seed; the paired design and small p-values make the ordering robust but do not quantify run-to-run variance. Within the conditioning, the $|f|$ encoding is part of every x–f configuration, and the mask spectrum and the DC residual were enabled together, so the individual shares of the three ingredients are not measured; the mask is supplied as its temporal spectrum because the regularizer

operates along temporal frequency, and we accept that this form of conditioning may not be the most intuitive one. All undersampling is retrospective, so effects specific to a prospective acquisition — motion, off-resonance, sequence constraints — are absent; a fresh mask each step keeps the network from fitting one pattern but does not substitute for prospective validation. Finally, we evaluate the four challenge metrics, not the derived quantities that motivate 4D flow MRI: they are computed on the velocity field inside the aortic mask, from which wall shear stress and flow volume follow, so they constrain those indirectly, but the transfer is not measured here.

Our components are drawn from existing work — the serial cascade skeleton [8,18], the CG data term [1], x–f sparsity [10] and feature-wise conditioning [13]; the contribution here is their integration for 4D flow MRI velocity reconstruction and a controlled measurement of what each one contributes.

Disclosure of Interests. The authors have no competing interests.

References

1. 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 (2019). https://doi.org/10.1109/TMI.2018.2865356
2. Baltes, C., Kozerke, S., Hansen, M.S., Pruessmann, K.P., Tsao, J., Boesiger, P.: Accelerating cine phase-contrast flow measurements using k-t BLAST and k-t SENSE. Magnetic Resonance in Medicine **54**(6), 1430–1438 (2005). https://doi.org/10.1002/mrm.20730
3. Bissell, M.M., et al.: 4D flow cardiovascular magnetic resonance consensus statement: 2023 update. Journal of Cardiovascular Magnetic Resonance **25**, 40 (2023). https://doi.org/10.1186/s12968-023-00942-z
4. Carlsson, M., Töger, J., Kanski, M., Markenroth Bloch, K., Ståhlberg, F., Heiberg, E., Arheden, H.: Quantification and visualization of cardiovascular 4D velocity mapping accelerated with parallel imaging or k-t BLAST: head to head comparison and validation at 1.5 T and 3 T. Journal of Cardiovascular Magnetic Resonance **13**(1), 55 (2011). https://doi.org/10.1186/1532-429X-13-55
5. CMRxRecon Challenge Organizers: CMRx4DFlow2026: Accelerated 4D flow MRI reconstruction challenge. https://cmrxrecon.github.io/2026/ (2026), MICCAI 2026 STACOM challenge. Accessed 28 July 2026
6. Dyverfeldt, P., et al.: 4D flow cardiovascular magnetic resonance consensus statement. Journal of Cardiovascular Magnetic Resonance **17**, 72 (2015). https://doi.org/10.1186/s12968-015-0174-5
7. Giese, D., Schaeffter, T., Kozerke, S.: Highly undersampled phase-contrast flow measurements using compartment-based k-t principal component analysis. Magnetic Resonance in Medicine **69**(2), 434–443 (2013). https://doi.org/10.1002/mrm.24273
8. Hammernik, K., Schlemper, J., Qin, C., Duan, J., Summers, R.M., Rueckert, D.: Systematic evaluation of iterative deep neural networks for fast parallel MRI reconstruction with sensitivity-weighted coil combination. Magnetic Resonance in Medicine **86**(4), 1859–1872 (2021). https://doi.org/10.1002/mrm.28827
9. Jacobs, L., Piccirelli, M., Vishnevskiy, V., Kozerke, S.: FlowMRI-Net: A generalizable self-supervised 4D flow MRI reconstruction network. Journal of Cardiovascular Magnetic Resonance **27**(2), 101913 (2025). https://doi.org/10.1016/j.jocmr.2025.101913
10. Jung, H., Sung, K., Nayak, K.S., Kim, E.Y., Ye, J.C.: k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI. Magnetic Resonance in Medicine **61**(1), 103–116 (2009). https://doi.org/10.1002/mrm.21757
11. Markl, M., Frydrychowicz, A., Kozerke, S., Hope, M., Wieben, O.: 4D flow MRI. Journal of Magnetic Resonance Imaging **36**(5), 1015–1036 (2012). https://doi.org/10.1002/jmri.23632
12. Pelc, N.J., Bernstein, M.A., Shimakawa, A., Glover, G.H.: Encoding strategies for three-direction phase-contrast MR imaging of flow. Journal of Magnetic Resonance Imaging **1**(4), 405–413 (1991). https://doi.org/10.1002/jmri.1880010404
13. Perez, E., Strub, F., de Vries, H., Dumoulin, V., Courville, A.: FiLM: Visual reasoning with a general conditioning layer. In: AAAI Conference on Artificial Intelligence. pp. 3942–3951 (2018). https://doi.org/10.1609/aaai.v32i1.11671
14. Qazi, S.A., Bianchessi, T., Viola, F., Trenti, C., Ylipää, E., Ebbers, T., Dyverfeldt, P.: FlowVN trained on a single dataset enables rapid reconstruction of highly accelerated 4D flow MRI across multiple sites. Magnetic Resonance in Medicine **96**(1), 374–386 (2026). https://doi.org/10.1002/mrm.70317

15. Qin, C., Duan, J., Hammernik, K., Schlemper, J., Küstner, T., Botnar, R., Prieto, C., Price, A.N., Hajnal, J.V., Rueckert, D.: Complementary time-frequency domain networks for dynamic parallel MR image reconstruction. Magnetic Resonance in Medicine **86**(6), 3274–3291 (2021). https://doi.org/10.1002/mrm.28917
16. Qin, C., Schlemper, J., Duan, J., Seegoolam, G., Price, A., Hajnal, J., Rueckert, D.: k-t NEXT: Dynamic MR image reconstruction exploiting spatio-temporal correlations. In: Medical Image Computing and Computer Assisted Intervention (MICCAI). LNCS, vol. 11765, pp. 505–513 (2019). https://doi.org/10.1007/978-3-030-32245-8_56
17. Sagawa, S., Koh, P.W., Hashimoto, T.B., Liang, P.: Distributionally robust neural networks for group shifts: On the importance of regularization for worst-case generalization. In: International Conference on Learning Representations (ICLR) (2020), arXiv:1911.08731
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 Transactions on Medical Imaging **37**(2), 491–503 (2018). https://doi.org/10.1109/TMI.2017.2760978
19. Tran, D., Wang, H., Torresani, L., Ray, J., LeCun, Y., Paluri, M.: A closer look at spatiotemporal convolutions for action recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 6450–6459 (2018). https://doi.org/10.1109/CVPR.2018.00675
20. Vishnevskiy, V., Walheim, J., Kozerke, S.: Deep variational network for rapid 4D flow MRI reconstruction. Nature Machine Intelligence **2**(4), 228–235 (2020). https://doi.org/10.1038/s42256-020-0165-6