Rethinking Real-World MRI Denoising: Learning from Physical Noise

Sebastian Rassmann[1], David Kügler[1], Sascha Brunheim[1], Philipp Ehses[1], Martin Reuter[1,2,3]

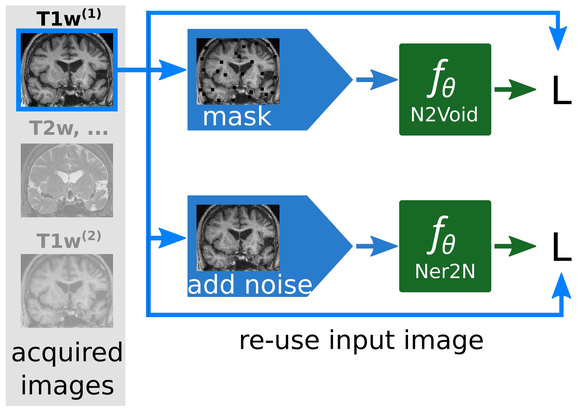
[1] DZNE [2] Martinos Center [3] Harvard Medical School — DeepMI Lab

ECCV 2026 | MALMÖ | SEPT 8–12

Noise in MRI

- MRI is inherently noisy ⇒ **no "clean" MRI exists**
  - approximation: average $N_{EX} \gg 1$ images
- Noise harms information extraction [1]
  - e.g. biomarkers like PVS and WMHs
- Canonical Noise2Clean training is impossible

Existing MRI Denoisers

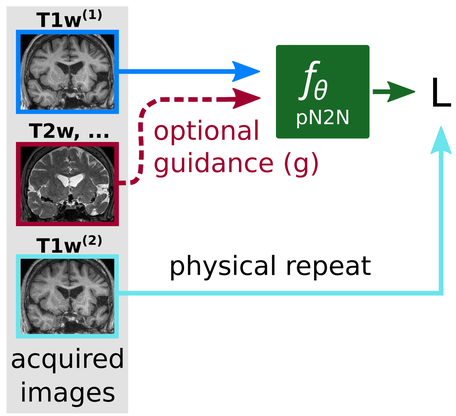


- Resort to surrogate single-noisy targets, e.g.
  - blind-spot masks: N2Void [2,3], Neigh2Neigh [4]
  - synthetic noising: Noisier2N [5], N2Score [6]

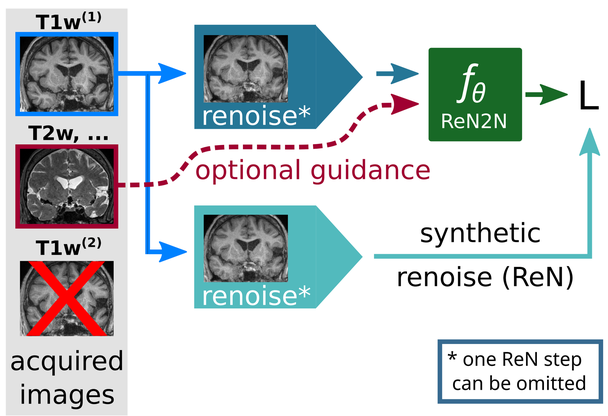
⇒ Tradeoff: **noise concealment vs. domain shift**
⇒ **incomplete noise removal and blurring**

Proposed method: YADO (You Accurately Denoise real Observations)

pN2N (physical Noise2Noise)**:**



ReN2N (ReNoise2Noise)**:**



- N2N-like [7] training **with / without guidance** using **physical / synthetic** (DM-generated) repeats

⇒ **4 YADO variants**: { **u** unguided, **g** guided } x { **pN2N**, **ReN2N** }

✔ **YADO: unprecedented image quality**

- **training: physical / realistic synthetic pairs**

⇒ **YADO + 6.5 (+5) min routine MRI ≈ 45 min acquisition**

✗ **Existing denoisers: disappointing performance**

- **training: surrogate target, not real-world MRI**

Quantitative Benchmark

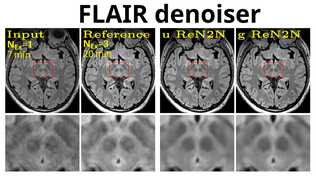
	OASIS-3 trained (ca. 1.0 mm)							HCP-A trained (0.8 mm)				all
	SSIM (%)						CNR	SSIM (%)			CNR	rank
Test dataset	OAS3	CHDI	Kirby	HCP-E	-YA	RS	PVS	HCP-E	-YA	RS	PVS	all
u pN2N	89.39	94.25	95.69	89.65	94.06	93.83	**2.63**	87.05	85.85	90.44	2.15	4.55
u ReN2N	89.41	94.33	**95.99**	90.60	94.50	94.37	**2.74**	88.25	86.66	91.23	**2.32**	2.09
g pN2N	**90.03**	94.11	(94.63)	89.96	94.11	**94.31**	**2.85**	86.87	86.05	89.93	2.25	4.17
g ReN2N	**89.77**	**94.59**	(95.63)	**90.91**	**94.59**	**94.59**	**2.85**	**88.74**	**86.98**	**91.69**	**2.29**	**1.46**
Ner2N [5]	88.59	93.53	95.58	89.47	93.91	93.42	2.28	84.63	82.65	88.45	2.10	9.26
N2Score [6]	89.07	93.77	95.21	87.89	93.61	92.75	1.06	79.97	65.71	90.24	1.71	12.24
N2V [8]	88.69	93.67	95.59	88.43	91.38	92.62	2.17	87.42	85.89	90.56	2.22	7.81
Neigh2N. [4]	87.94	93.03	94.59	89.08	93.61	92.68	2.31	87.36	85.40	89.63	1.83	9.75
CNN3D [9]	87.79	93.00	95.11	88.86	92.62	92.62	2.25	86.28	84.51	89.31	1.95	11.62
DDM² [10,1]	87.28	91.14	92.73	84.10	89.80	91.28	1.76	72.24	75.28	84.41	1.01	17.70
BME-X [11]	83.10	87.81	86.64	82.87	88.38	88.05	1.55	82.71	82.17	88.14	1.56	18.68
N2Contr [12]	87.32	93.12	95.34	87.98	90.46	92.16	2.42	85.51	83.39	88.97	2.12	11.82
DIP [13]	87.87	93.27	95.31	89.04	92.23	92.88	2.33	86.77	84.86	89.58	1.97	9.84
BM4D [14]	88.28	93.40	95.33	89.37	93.71	92.93	2.13	87.81	86.09	90.27	1.83	7.96
Identity	86.91	92.66	95.03	87.09	89.50	91.60	2.33	84.61	82.68	88.47	2.10	14.35

bold: best ($p<0.05$) — gray: no improvement over Identity (NOP, $p \geq 0.05$)

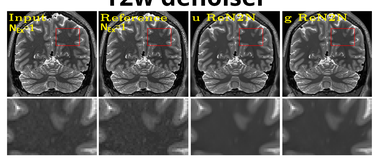
⇒ **All YADO versions achieve SOTA performance**
- 7/17 baseline worse than Identity (MedFM poster)
- Only 2 competing methods actually improve CNR at PVS

Other contrasts & OOD

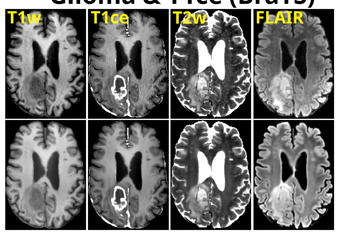
FLAIR denoiser



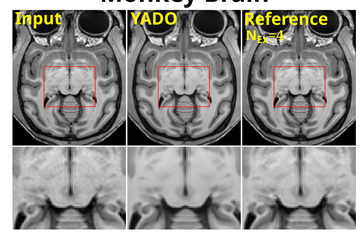
T2w denoiser



Glioma & T1ce (BraTS)

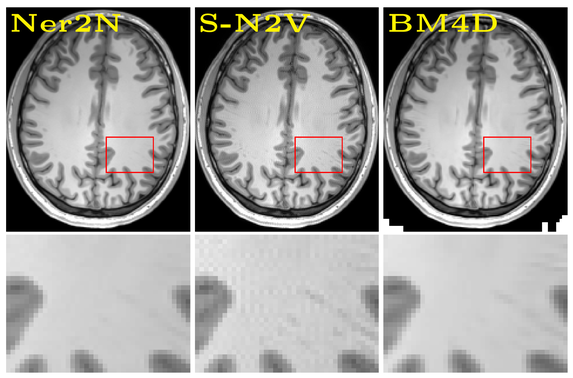


Monkey Brain

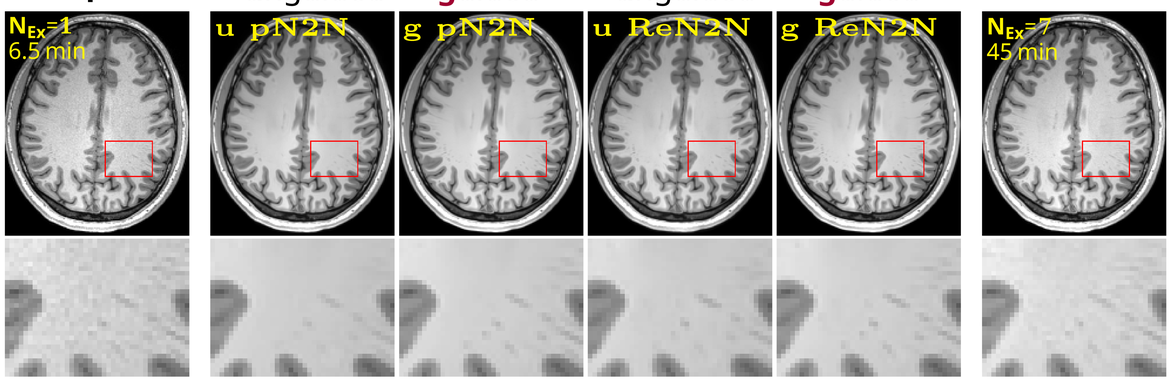


Previous SOTA

Existing Denoisers



Our Results



Conclusion

- Previous methods fail on real-world noise
  - ⇒ bad performance hidden by flawed evaluation
- Our work closes this gap:

SOTA denoising from routine scans

⇒ reliable image quality improvements

⇒ matches physical noise supression

code & weights

References: [1] Rassmann et al. "Regression is all you need for Medical Image Translation", IEEE TMI '26; [2] Krull et al., "Noise2Void: Learning Denoising from Single Noisy Images", CVPR '19; [3] Batson et al., "Noise2Self: Blind Denoising by Self-Supervision", ICML '19; [4] Huang et al., "Neighbor2Neighbor: Self-Supervised Denoising from Single Noisy Images", CVPR '21; [5] Moran et al., "Noisier2Noise: Learning to Denoise from Unpaired Noisy Data", CVPR '20; [6] Kim et al., "Noise2Score: Tweedie's Approach to Self-Supervised Image Denoising", NeurIPS '21; [7] Lehtinen et al., "Noise2Noise: Learning Image Restoration without Clean Data", ICML '18; [8] Broaddus et al., "Removing Structured Noise with Self-Supervised Blind-Spot Networks", ISBI '20; [9] Ran et al., "Denoising of 3D MRI Using a Residual Encoder-Decoder wGAN", MedIA '19; [10] Xiang et al., "DDM2: Self-Supervised Diffusion MRI Denoising with Generative Diffusion Models", ICLR '23; [11] Sun et al., "A Foundation Model for Enhancing MRI and Downstream Segmentation, Registration and Diagnostic Tasks", Nat Biomed Eng '25; [12] Wagner et al., "Noise2Contrast: Multi-Contrast Fusion Enables Self-Supervised Tomographic Image Denoising", IPMI '23; [13] Ulyanov et al., "Deep Image Prior DIP", CVPR '18; [14] Maggioni et al., "Nonlocal Transform-Domain Filter for Volumetric Data Denoising and Reconstruction", IEEE TIP '12