Supplemenatry Materials of MS-Glance: Bio-Inspired Non-semantic Context Vectors and their Applications in Supervising Image Reconstruction

Here, we show some additional results mentioned in the main paper: qualitative results on fitting the Astronaut image and ablation studies on the Glance's window kernel and distance measure. We also add additional details on MS-Glance's implementation and the network architecture of DRDN, which we use for undersampled MRI reconstruction experiments.

1. More qualitative results of *Astronaut*

Astronaut is a color image of the astronaut Eileen Collins. In Figure 1, we compare the step-wise reconstruction of *Astronaut* by SIREN and SIREN+MSGlance. The reconstructed images and the corresponding SSIM error maps are visualized. MS-Glance reconstructs the image details faster (the blue boxes in step 40) and ends up with a finer reconstruction (the blue boxes in step 500).

2. More ablation studies

2.1. Uniform Kernel and Gaussian Kernel

The uniform window kernel is a key distinction between the Glance Index Measure and methods like SSIM and S3IM. To conduct a comprehensive ablation study, we replace our uniform kernel with their Gaussian kernel on both tasks. For MRI reconstruction, we use the IXI dataset under two acceleration rates. For INR fitting, we use the Coco dataset.

Table 1 states that compared with the Gaussian kernel, our uniform kernel not only stabilizes training but also enhances performance. Initially, when we applied the Gaussian kernel to MS-Glance, it caused the loss function to produce NaN values. To mitigate this, we detected NaN and switched to a standard L_p loss during NaN iterations. However, a significant number of steps still resulted in NaN values. To further investigate, we decomposed MS-Glance. We observed that both Local Glance with Gaussian and MS-Glance with Gaussian led to approximately 60% NaN loss, contributing to the large performance degradation. While the Global Glance with Gaussian's training remained stable, it also experienced a performance decline.

	Undersampled MRI reconstruction 5x		7x		INR fitting	
	PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM
MS-Glance	**31.434**	**0.9535**	**30.865**	**0.9485**	**35.249**	**0.9493**
MS-Glance - Gaussian	29.497	0.9323	28.699	0.9229	35.099	0.9469
Global Glance	**31.122**	**0.9524**	**30.711**	**0.9485**	**34.843**	**0.9439**
Global Glance - Gaussian	31.104	0.9524	30.622	0.9483	34.827	0.9441
Local Glance	**31.346**	**0.9537**	**30.813**	**0.9483**	**35.004**	**0.9463**
Local Glance - Gaussian	29.487	0.9315	28.573	0.9208	34.830	0.9430

Table 1. Ablation study of the window kernel.

2.2. Glance Index Measure and SSIM

While we compare our method with SSIM loss in all experiments, we also highlight the connection between the Glance Index Measure and SSIM, which is discussed in detail in the main paper. In this section, we provide additional experimental results to compare the performance of the Glance Index Measure against SSIM. As mentioned in the main paper, the structural term of SSIM computes covariance similarly to how the Glance Index Measure operates. However, SSIM also incorporates luminance (l) and contrast (c) terms. To account for this, we extend our Glance Index Measure by integrating the computation of l and c, multiplying them with the original Glance Index Measure. We tested this modified approach across both tasks.

We perform the evaluation on both tasks. For MRI reconstruction, we use the IXI dataset under two acceleration rates. For INR fitting, we use the Coco dataset. Table 2 demonstrates the effectiveness of the Glance Index Measure, particularly in global scenarios. The current Glance Index Measure shows that MS-Glance and Global Glance remain superior. However, the Local Glance enhanced with l and c exhibits improved performance, especially in SSIM computations. This improvement is expected, as it directly optimizes a term similar to SSIM itself. Additionally, we explored combining the original Global Glance design with the new Local Glance incorporating l and c, with results shown in the last row. This approach, however, did not perform as well as the original MS-Glance design, suggesting a conflict between the two approaches.

Figure 1. Step-wise reconstruction of the example image, *Astronaut*

	Undersampled MRI reconstruction 5x		Undersampled MRI reconstruction 7x		INR fitting	
	PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM
MS-Glance	**31.434**	**0.9535**	**30.865**	**0.9485**	**35.249**	**0.9493**
+ l and c of SSIM	31.035	0.9523	30.660	0.9483	35.131	0.9484
Global Glance (a)	**31.122**	**0.9524**	**30.711**	**0.9485**	**34.843**	0.9439
+ l and c of SSIM	31.055	0.9523	30.653	0.9479	34.714	**0.9426**
Local Glance	31.346	0.9537	30.813	0.9483	**35.004**	0.9463
+ l and c of SSIM (b)	**31.425**	**0.9557**	**30.939**	**0.9519**	34.913	**0.9468**
Combination of (a) and (b)	30.953	0.9516	30.701	0.9492	35.242	0.9496

Table 2. The effect of Glance Index Measure and SSIM to MS-Glance.

3. Additional Details

3.1. Implementation of MS-Glance

In the Global Glance process, we randomly select pixels and shuffle them 10 times, resulting in more Glance vectors for computing the Global Glance Index Measure. As shown in Table 3, shuffling leads to a slight improvement in performance. The experiments are carried out on the Coco dataset.

Shuffle times	1	5	10
PSNR	35.202	**35.257**	35.249
SSIM	0.9489	0.9491	**0.9493**

Table 3. The effect of the shuffle time on INR fitting.

3.2. Architecture of DRDN

We choose DRDN as the network for undersampled MRI reconstruction. Its strong performance has been validated by their original experiments and many recently established works [1, 2]. DRDN [3] customizes the local and global structure design for the MRI reconstruction task. It uses a Squeeze-and-excitation Dilated Residual Dense Block (SDRDB) as the backbone. The main diagram is shown in Figure 2.

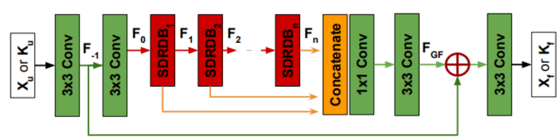

Figure 2. The diagram of DRDN [3].

Globally, DRDN consists of an initial feature extraction module (two sequential 3×3 convolution layers), multiple SDRDBs followed by global feature fusion (a concatenation operation for all SDRDBs' output), and global residual learning enhanced by a Squeeze-and-excitation on the residual branches.

The structure in each SDRDB is shown in Figure. In each SDRDB, there are four densely connected atrous convolution layers, local feature fusion, Squeeze-and-Excitation, and local residual learning.

References

[1] Hyungjin Chung and Jong Chul Ye. Score-based diffusion models for accelerated mri. *Medical Image Analysis*, page 102479, 2022. 2

[2] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations.

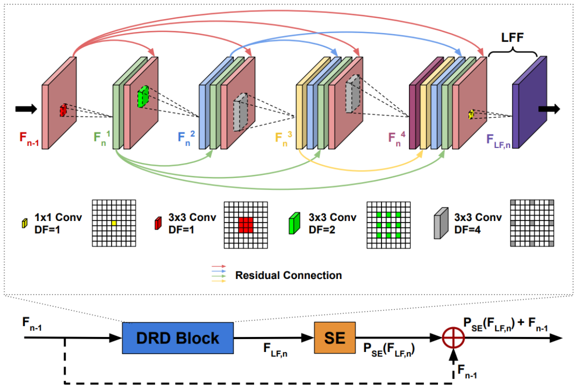

Figure 3. The component of each SDRDB.

In *International Conference on Learning Representations*, 2021. 2

[3] Bo Zhou and S Kevin Zhou. Dudornet: learning a dual-domain recurrent network for fast mri reconstruction with deep t1 prior. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4273–4282, 2020. 2