# DC-WCNN: A DEEP CASCADE OF WAVELET BASED CONVOLUTIONAL NEURAL NETWORKS FOR MR IMAGE RECONSTRUCTION

**Sriprabha Ramanarayanan**, Balamurali Murugesan, Keerthi Ram, Mohanasankar Sivaprakasam

Healthcare Technology Innovation Centre
Indian Institute of Technology Madras, India

April 6, 2020

# Magnetic Resonance (MR) Image Reconstruction

- MR Imaging modality - non-invasive, no ionizing radiation, provides high resolution images with excellent soft tissue contrast
- Long acquisition times  leads to patient discomfort
- Sub-Nyquist sampling of k-space done to accelerate acquisition. Image quality compromised. Resulting image has aliasing artifacts.
- Mapping undersampled k-space to fully sampled image is done using Compressed Sensing (CS-MRI).

# CS-MRI Reconstruction

- Acquired MRI images are compressible. This implicit sparsity in image or a suitable transform domain is used as prior.
- The aliasing artifacts due to k-space undersampling be incoherent (noise like) in that transform domain
- A non-linear reconstruction to enforce both sparsity of the image representation and consistency with the acquired data.

# Deep learning based MRI Reconstruction

- Convolutional Neural networks (CNNs) learn an end to end mapping between the undersampled (US) and fullysampled (FS) image in a data driven manner
- U-Net (Ronneberger et al), an encoder-decoder CNN based architecture, has shown to provide good results in MRI Reconstruction.
- Deep cascaded architectures closely mimic the iterative reconstruction of CS-MRI and have shown to give promising results. (DC-CNN by Schlemper et al. 2017).

# MRI Reconstruction using U-Net

- The latent space of the bottleneck layer of U-Net learns the non-linear compressed sensing representation (Hyun et al [1], Sun et al [2]).
- The hierarchical scale levels due to convolution and pooling layers provide good receptive field (Liu et al [3]).

[1] Chang Min Hyun, Hwa Pyung Kim, Sung Min Lee, Sungchul Lee, and Jin Keun Seo, “Deep learning for undersampled MRI reconstruction,” Physics in Medicine & Biology, vol. 63, no. 13, pp. 135007, jun 2018.

[2] Liyan Sun, Zhiwen Fan, Xinghao Ding, Yue Huang, and John Paisley, “Joint cs-mri reconstruction and segmentation with a unified deep network,” in Information Processing in Medical Imaging, 2019, pp. 492–504.

[3] Pengju Liu, Hongzhi Zhang, Kai Zhang, Liang Lin, and Wangmeng Zuo, “Multi-level wavelet-cnn for image restoration,” in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, June 2018.

# MRI Reconstruction using U-Net - Problems

- Sub-sampling done in the pooling layers in U-Net causes information loss.
- A fully convolutional network or Dilated convolutions are not better choices.
- Hence no compromise in receptive field, signal loss or the hierarchical architecture be made.

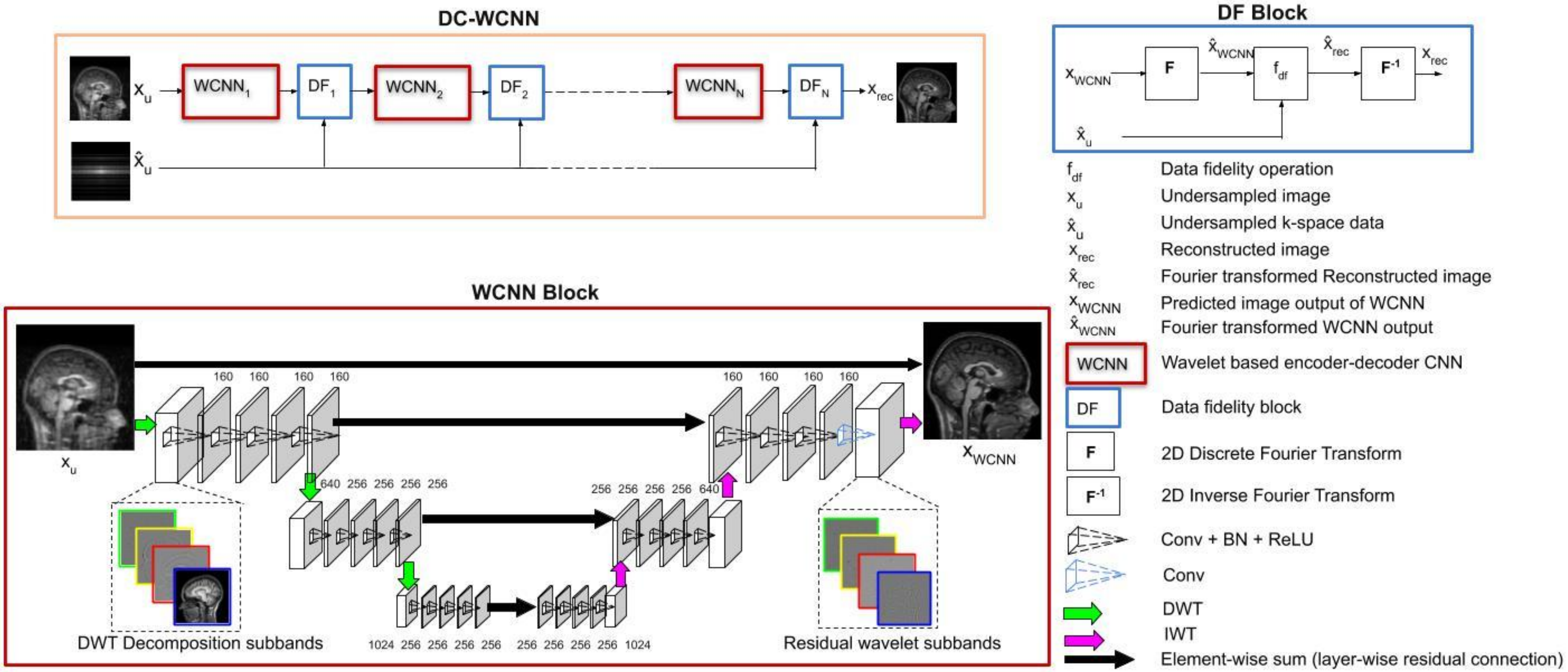
# MRI Reconstruction using U-Net - Proposed modification

- Replace the pooling layers with wavelet transform based subsampling layers.
- We use multi-level Haar wavelet packet transform - 1. Frequency and localization characteristics 2. Lossless partitioning of feature maps into orthogonal subbands 3. Sparsity.
- Resulting architecture - Wavelet CNN and Deep cascade Wavelet CNN

# Contributions

- We propose a wavelet based convolutional encoder - decoder neural network WCNN, for MR image reconstruction, with better signal representation. The proposed WCNN has residual connections, wavelet decomposition and recomposition operations in place of pooling and unpooling layers respectively.
- We also propose a deep cascaded architecture called DC-WCNN, by cascading a series of WCNN and data fidelity (DF) units for the MRI reconstruction problem.

# Proposed Architecture - DC-WCNN



# Problem formulation - MR Reconstruction

The proposed method using WCNN can be formulated by minimizing the optimization problem

$$\underset{x,\theta}{\operatorname{argmin}} ||x - f_{wcnn}(x_u|\theta)||_2^2 + \alpha||F_u x - y||_2^2$$

| | |
|---|---|
| $x \in C^N$ | the desired image to be reconstructed |
| $y \in C^M,\ y = F_u x$ | undersampled k-space measurements $M << N$ |
| $x_u = F_u^H y$ | zero-filled reconstruction |
| $x_{wcnn} = f_{wcnn}(x_u\|\theta)$ | $f_{wcnn}$ is the forward mapping of WCNN parameterized by $\theta$ |
| | $F_u$ is the undersampled Fourier encoding matrix |

# Problem formulation - Data Fidelity

- Data fidelity (DF) unit in k-space domain to ensure consistency with the acquired k-space measurements

$$\hat{x}_{rec} = \begin{cases} \hat{x}_{wcnn}(k) & k \notin \Omega \\ \frac{\hat{x}_{wcnn}(k)+\lambda \hat{x}_u(k)}{1+\lambda} & k \in \Omega \end{cases}$$

$\hat{x}_{wcnn} = F_f x_{wcnn}, \hat{x}_u = F_f x_u$ $\quad F_f$ is the Fourier encoding matrix

$\Omega$ is the index set of known k-space data

$\hat{x}_{rec}$ is the corrected k-space and $\lambda \to \infty$

# Problem formulation - Deep Cascade

- The proposed cascaded architecture, DC-WCNN, is a series of N such WCNN and DF units which can be formulated as,

$$x_n = WCNN_n(x_{n-1}^{df}) + x_{n-1}^{df}$$
$$x_n^{df} = DF_n(x_n)$$

$WCNN_n$ and $DF_n$ denote the $n^{th}$ WCNN reconstruction block and DF unit respectively, $n = 1, 2..N$
$x_0^{df} = x_u$ is the input to the first WCNN unit.
$x_{rec} = x_N^{df}$ is the output of the last DF unit.

# Problem formulation - DWT and IWT layers

- The 2D Discrete Wavelet Transform (DWT) and inverse wavelet transform (IWT) layers of WCNN are given by,

$$(X_1, X_2, ..., X_K) = DWT(X_{in})$$
$$X = IWT(X_1, X_2, ..., X_K)$$

$X_{in}$ the input image or an intermediate feature map
K is the number of subbands K = 4
with approximation, horizontal, vertical and diagonal subband
$X_k, k = 1, 2, ...K$ denote the subband coefficients
X is output of wavelet recomposition.

# Proposed architecture

- The DWT subbands are stacked and passed to the convolutional layers. The subbands are jointly learnt. Hence inter-dependencies of feature maps are learnt, enhancing spatial context.
- Residual connections - feature maps from contracting paths are element-wise added to the respective expanding path feature maps.
- DC-WCNN is similar to unfolding the optimization process of CSMRI and improves the performance of MRI reconstruction.

# Experiments and Results

- Dataset - Kirby 21 human brain dataset
- The dataset consists of 5460 slices of size 256x256 taken from 42 T1- weighted MPRAGE volumes out of which, 3770 slices from 29 volumes are used for training and 1690 slices from 13 volumes for validation.
- Fixed cartesian mask, 5x acceleration
- L2 loss function
- Adam optimizer, learning rate - 0.001
- Standalone mode and Deep cascaded mode

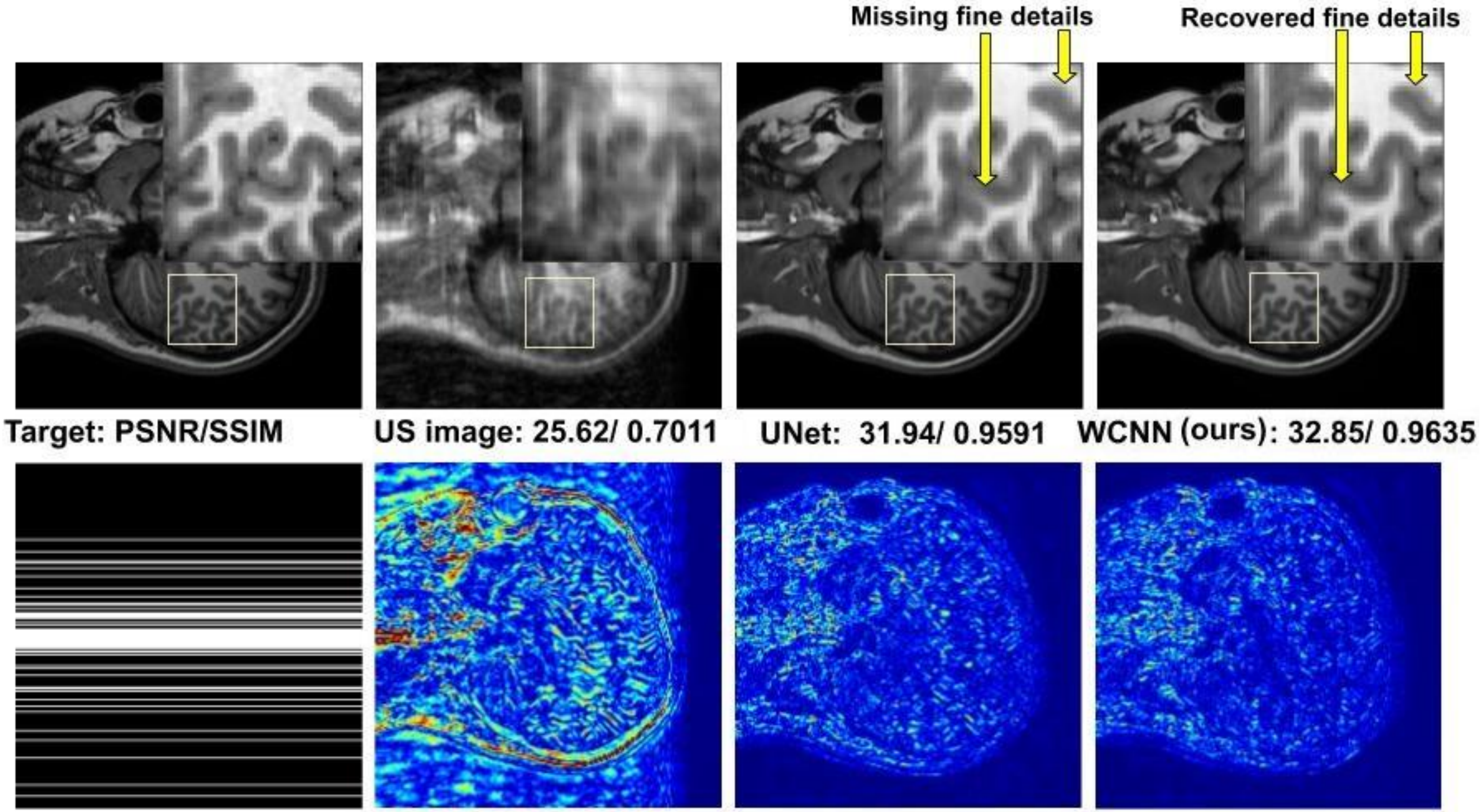
# Results and Discussions - Quantitative

**Table 1**. PSNR, SSIM, NMSE and HFEN results for 5x undersampling

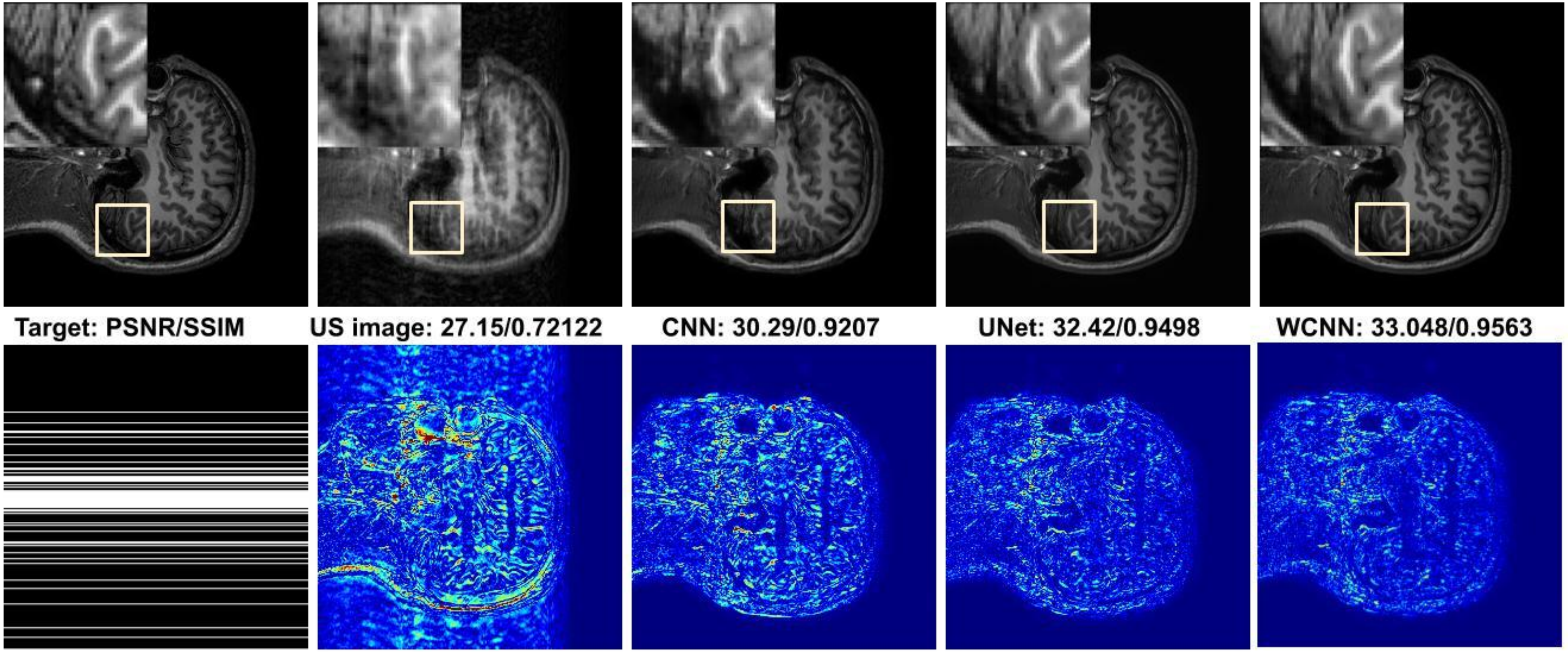
| | No. of cascades | Model | NMSE | PSNR | SSIM | HFEN |
|---|---|---|---|---|---|---|
| | - | US image | 0.07014 +/- 0.01 | 25.95 +/- 1.29 | 0.6056 +/- 0.05 | 0.7571 +/- 0.01 |
| Standalone mode | - | CNN | 0.02775 +/- 0.00 | 29.97 +/- 1.33 | 0.8897 +/- 0.02 | 0.5311 +/- 0.07 |
| | | U-Net | 0.01929 +/- 0.01 | 31.57 +/- 1.17 | 0.9302 +/- 0.01 | 0.4583 +/- 0.06 |
| | | U-NetMean | 0.01955 +/- 0.00 | 31.51 +/- 1.21 | 0.9303 +/- 0.01 | 0.4574 +/- 0.06 |
| | | WCNN (Ours) | **0.01466 +/- 0.00** | **32.75 +/- 1.55** | **0.9361 +/- 0.02** | **0.4228 +/- 0.06** |
| Deep cascade mode | 1 | DC-CNN | 0.0179 +/- 0.00 | 31.88 +/- 1.50 | 0.8784 +/- 0.02 | 0.4776 +/- 0.07 |
| | | DC-UNet | 0.01181 +/- 0.00 | 33.69 +/- 1.56 | 0.9294 +/- 0.02 | 0.4033 +/- 0.06 |
| | | DC-WCNN (Ours) | **0.01151 +/- 0.00** | **33.8 +/- 1.61** | **0.9308 +/- 0.02** | **0.3879 +/- 0.06** |
| | 2 | DC-CNN | 0.01249 +/- 0.00 | 33.45 +/- 1.69 | 0.9385 +/- 0.02 | 0.4229 +/- 0.07 |
| | | DC-UNet | 0.00884 +/- 0.00 | 34.95 +/- 1.70 | 0.9587 +/- 0.01 | 0.357 +/- 0.064 |
| | | DC-WCNN (Ours) | **0.00811 +/- 0.00** | **35.33 +/- 1.83** | **0.9614 +/- 0.01** | **0.3458 +/- 0.066** |
| | 3 | DC-CNN | 0.01029 +/- 0.00 | 34.3 +/- 1.82 | 0.9526 +/- 0.02 | 0.3912 +/- 0.07 |
| | | DC-UNet | 0.0077 +/- 0.00 | 35.56 +/- 1.77 | 0.9652 +/- 0.01 | 0.3378 +/- 0.07 |
| | | DC-WCNN (Ours) | **0.00682 +/- 0.00** | **36.09 +/- 1.96** | **0.9685 +/- 0.01** | **0.3236 +/- 0.06** |

WCNN has only 3 subsampling levels while UNet and UNetMean has four subsampling layers

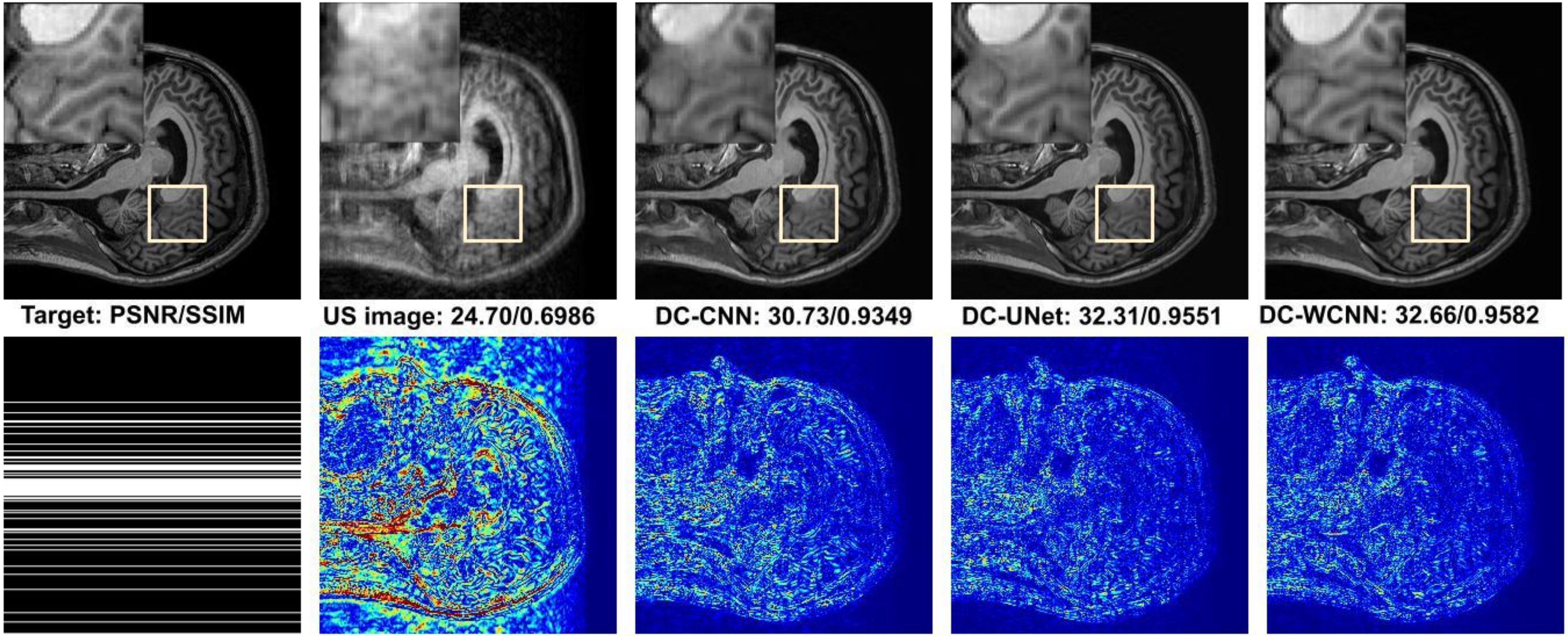
# Results and Discussions - Qualitative



# Qualitative Results for standalone mode



# Qualitative Results for deep cascade mode



# Conclusions

- This work presents a combination of successful architectures - Deep cascaded CNN and encoder-decoder based CNNs addressing the problem of information loss in pooling/unpooling layers.
- The WCNN is a variant of U-Net with DWT in place of pooling and IWT in place of unpooling. The DC-WCNN architecture is a cascade of WCNN and DF units.
- Experimental results show that WCNN and DC-WCNN outperform other compared methods for 5x acceleration factor.
- Better recovery of fine structures.

# Thank you

Paper

https://arxiv.org/abs/2001.02397

Code

https://github.com/sriprabhar/DC-WCNN

Contact

sriprabha.r@htic.iitm.ac.in