## I. Proof of Lemma 1

**Lemma 1** If $\boldsymbol{W}_S$ has an upper bound, then we have

$$||\boldsymbol{W}_S\boldsymbol{W}_A\boldsymbol{x} - \boldsymbol{W}_S Soft(\boldsymbol{W}_A\boldsymbol{x}, \boldsymbol{b})||_2^2 \leq c\sigma^2 \tag{1}$$

for some universal constant $c$ independent of $\sigma$.

Proof: From Eqn. (1), we have

$$\begin{aligned} &||\boldsymbol{W}_S\boldsymbol{W}_A\boldsymbol{x} - \boldsymbol{W}_S Soft(\boldsymbol{W}_A\boldsymbol{x}, \boldsymbol{b})||_2^2 \\ &\leq ||\boldsymbol{W}_S||_2^2 \, ||\boldsymbol{W}_A\boldsymbol{x} - Soft(\boldsymbol{W}_A\boldsymbol{x}, \boldsymbol{b})||_2^2 . \end{aligned} \tag{2}$$

A convolutional layer with RealSN constraint [1] is utilized to generate the synthesis sparsifying transform, which makes $\|\boldsymbol{W}_S\|_2^2$ have an upper bound. Defined the upper bound is $\alpha$, and the upper bound of Eqn. (2) can be further determined by

$$\begin{aligned} &||\boldsymbol{W}_S\boldsymbol{W}_A\boldsymbol{x} - \boldsymbol{W}_S Soft(\boldsymbol{W}_A\boldsymbol{x}, \boldsymbol{b})||_2^2 \\ &\leq \alpha \, ||\boldsymbol{W}_A\boldsymbol{x} - Soft(\boldsymbol{W}_A\boldsymbol{x}, \boldsymbol{b})||_2^2 . \end{aligned} \tag{3}$$

Let $(\boldsymbol{W}_A\boldsymbol{x})_i$ represent the $i$-th element of $\boldsymbol{W}_A\boldsymbol{x}$ and $\beta_i$ denote the $i$-th element of $\boldsymbol{b}$. The soft thresholding operator $Soft[(\boldsymbol{W}_A\boldsymbol{x})_i, \beta_i]$ is defined as

$$Soft[(\boldsymbol{W}_A\boldsymbol{x})_i, \beta_i] = \begin{cases} (\boldsymbol{W}_A\boldsymbol{x})_i + \beta_i, & (\boldsymbol{W}_A\boldsymbol{x})_i < -\beta_i \\ 0, & |(\boldsymbol{W}_A\boldsymbol{x})_i| \leq \beta_i \\ (\boldsymbol{W}_A\boldsymbol{x})_i - \beta_i, & (\boldsymbol{W}_A\boldsymbol{x})_i > \beta_i, \end{cases} \tag{4}$$

then one of the following situations will happen.

**CASE 1:** Any element of $\boldsymbol{W}_A\boldsymbol{x}$ satisfies $(\boldsymbol{W}_A\boldsymbol{x})_i < -\beta_i$, then $Soft(\boldsymbol{W}_A\boldsymbol{x}, \boldsymbol{b}) = \boldsymbol{W}_A\boldsymbol{x} + \boldsymbol{b}$. Therefore, we have

$$\begin{aligned} &||\boldsymbol{W}_S\boldsymbol{W}_A\boldsymbol{x} - \boldsymbol{W}_S Soft(\boldsymbol{W}_A\boldsymbol{x}, \boldsymbol{b})||_2^2 \\ &\leq \alpha \, ||\boldsymbol{W}_A\boldsymbol{x} - \boldsymbol{W}_A\boldsymbol{x} - \boldsymbol{b}||_2^2 \\ &= \alpha \, ||\boldsymbol{b}||_2^2 \end{aligned} \tag{5}$$

**CASE 2:** Any element of $\boldsymbol{W}_A\boldsymbol{x}$ satisfies $|(\boldsymbol{W}_A\boldsymbol{x})_i| \leq \beta_i$, then $Soft(\boldsymbol{W}_A\boldsymbol{x}, \boldsymbol{b}) = 0$. Therefore, we have

$$\begin{aligned} &||\boldsymbol{W}_S\boldsymbol{W}_A\boldsymbol{x} - \boldsymbol{W}_S Soft(\boldsymbol{W}_A\boldsymbol{x}, \boldsymbol{b})||_2^2 \\ &\leq \alpha \, ||\boldsymbol{W}_A\boldsymbol{x}||_2^2 . \end{aligned} \tag{6}$$

Since $|(\boldsymbol{W}_A\boldsymbol{x})_i| \leq \beta_i$, then $||\boldsymbol{W}_A\boldsymbol{x}||_2^2 \leq ||\boldsymbol{b}||_2^2$, it follows that

$$\begin{aligned} &||\boldsymbol{W}_S\boldsymbol{W}_A\boldsymbol{x} - \boldsymbol{W}_S Soft(\boldsymbol{W}_A\boldsymbol{x}, \boldsymbol{b})||_2^2 \\ &\leq \alpha \, ||\boldsymbol{b}||_2^2 . \end{aligned} \tag{7}$$

**CASE 3:** Any element of $\boldsymbol{W}_A\boldsymbol{x}$ satisfies $(\boldsymbol{W}_A\boldsymbol{x})_i > \beta_i$, then $Soft(\boldsymbol{W}_A\boldsymbol{x}, \boldsymbol{b}) = \boldsymbol{W}_A\boldsymbol{x} - \boldsymbol{b}$. Therefore, we have

$$\begin{aligned} &||\boldsymbol{W}_S\boldsymbol{W}_A\boldsymbol{x} - \boldsymbol{W}_S Soft(\boldsymbol{W}_A\boldsymbol{x}, \boldsymbol{b})||_2^2 \\ &\leq \alpha \, ||\boldsymbol{W}_A\boldsymbol{x} - \boldsymbol{W}_A\boldsymbol{x} + \boldsymbol{b}||_2^2 \\ &= \alpha \, ||\boldsymbol{b}||_2^2 . \end{aligned} \tag{8}$$

**CASE 4: CASE 4** is a union of the **CASE 1**, **CASE 2** and **CASE 3**. Therefore, as long as we find the upper bound of $||\boldsymbol{W}_S\boldsymbol{W}_A\boldsymbol{x} - \boldsymbol{W}_S Soft(\boldsymbol{W}_A\boldsymbol{x}, \boldsymbol{b})||_2^2$ under the **CASE 1-3**, the upper bound of that under the **CASE 4** will be determined. Hence, when **CASE 4** happens, the upper bound of $||\boldsymbol{W}_S\boldsymbol{W}_A\boldsymbol{x} - \boldsymbol{W}_S Soft(\boldsymbol{W}_A\boldsymbol{x}, \boldsymbol{b})||_2^2$ is the maximum upper bound of the first three situations. Based on Eqn. (5), Eqn. (7) and Eqn. (8), we have

$$||\boldsymbol{W}_S\boldsymbol{W}_A\boldsymbol{x} - \boldsymbol{W}_S Soft(\boldsymbol{W}_A\boldsymbol{x}, \boldsymbol{b})||_2^2 \leq \alpha \, ||\boldsymbol{b}||_2^2 \tag{9}$$

Recall the definition of threshold vectors $\boldsymbol{b} = \boldsymbol{c} \otimes \boldsymbol{s}$, and each element of proportional constant vectors $\boldsymbol{c}$ has a limited range $c_i \in [c_{min}, c_{max}]$. Let $\beta_{max} = c_{max} \cdot \sigma$ denote the maximum element of $\boldsymbol{b}$. Thus,

$$||\boldsymbol{W}_S\boldsymbol{W}_A\boldsymbol{x} - \boldsymbol{W}_S Soft(\boldsymbol{W}_A\boldsymbol{x}, \boldsymbol{b})||_2^2 \leq \alpha c_{max}^2 \sigma^2 \tag{10}$$

## II. Proof of Lemma 2

**Lemma 2** (The bounded gradient of data fidelity function) Consider the sampling model $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}$ and $\boldsymbol{y} = \boldsymbol{F}_u\boldsymbol{x} + \boldsymbol{n}$, where $||\boldsymbol{n}||_2 \leq \epsilon$, then the functions $f(\boldsymbol{x}) = \frac{1}{2}||\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}||_2^2$ and $f(\boldsymbol{x}) = \frac{1}{2}||\boldsymbol{y} - \boldsymbol{F}_u\boldsymbol{x}||_2^2$ have the bounded gradient, respectively.

Proof:

(In the SCI task) The gradient of $f(\boldsymbol{x}) = \frac{1}{2}||\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}||_2^2$ in the SCI formulation is

$$\nabla f(\boldsymbol{x}) = \boldsymbol{H}^\top(\boldsymbol{H}\boldsymbol{x} - \boldsymbol{y}) \tag{11}$$

thus, we have

$$\begin{aligned} \|\nabla f(\boldsymbol{x})\|_2 &\leq \left\|\boldsymbol{H}^\top\boldsymbol{H}\boldsymbol{x}\right\|_2 + \left\|\boldsymbol{H}^\top\boldsymbol{y}\right\|_2 \\ &\leq \left\|\boldsymbol{H}^\top\right\|_2 \left(\|\boldsymbol{H}\boldsymbol{x}\|_2 + \|\boldsymbol{y}\|_2\right) . \end{aligned} \tag{12}$$

Since the values in the sensing matrix $\boldsymbol{H}$ and the pixel values in the image $\boldsymbol{x}$ are normolized into $[0, 1]$, we obtain

$$\|\boldsymbol{H}\boldsymbol{x}\|_2 \leq \|\boldsymbol{H}\|_2 \, \|\boldsymbol{x}\|_2 \leq \sqrt{\lambda_1 N}. \tag{13}$$

where $\lambda_1$ denotes the maximum eigenvalue of $\boldsymbol{H}^\top\boldsymbol{H}$, and $N$ represents the pixel values of the image. In Eqn. (13), $\|\boldsymbol{y}\|_2 = \|\boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}\|_2$ and $\|\boldsymbol{n}\|_2 \leq \epsilon$, here, $\boldsymbol{x}$ is the image whose pixel values are normalized into $[0, 1]$. Therefore,

$$\|\boldsymbol{y}\|_2 = \|\boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}\|_2 \leq \|\boldsymbol{H}\boldsymbol{x}\|_2 + \|\boldsymbol{n}\|_2 \leq \sqrt{\lambda_1 N} + \epsilon. \tag{14}$$

Let $\lambda_2$ denote the maximum eigenvalue of $\boldsymbol{H}\boldsymbol{H}^\top$, then we have

$$\left\|\boldsymbol{H}^\top\right\|_2 \leq \sqrt{\lambda_2}. \tag{15}$$

Substituting Eqn. (13) and Eqn. (14) Eqn. (15) into Eqn. (12), we get

$$\begin{aligned} \|\nabla f(\boldsymbol{x})\|_2 &\leq \left\|\boldsymbol{H}^\top\right\|_2 \left(\|\boldsymbol{H}\boldsymbol{x}\|_2 + \|\boldsymbol{y}\|_2\right) \\ &\leq \sqrt{\lambda_2}\left(2\sqrt{\lambda_1 N} + \epsilon\right) . \end{aligned} \tag{16}$$

This completes the proof of **Lemma 2** in the SCI task.

(In the CSMRI task) See [2] for detailed proof.

## III. Proof of Theorem 2

**Theorem 2** (Convergence of the proposed algorithms) Based on **Lemma 2** and **Theorem 1**, as well as the value of $\sigma$ satisfies the condition of sequentially diminishing as the unfolding stage $t$ increases in the proposed DUN-BSTNet algorithm, then both $\boldsymbol{x}^t$ and $\boldsymbol{v}^t$ will converge, respectively.

Proof: Following [3], we define $\boldsymbol{\theta}^t = (\boldsymbol{x}^t, \boldsymbol{v}^t)$, and let $\boldsymbol{\Theta}$ be the domain of $\boldsymbol{\theta}^t$ for all stages $t$. On the domain $\boldsymbol{\Theta}$, we define a distance function $D$: $\boldsymbol{\Theta} \times \boldsymbol{\Theta} \to \mathbb{R}$ such that

$$D(\boldsymbol{\theta}^t, \boldsymbol{\theta}^{t-1}) = \left\|\boldsymbol{x}^t - \boldsymbol{x}^{t-1}\right\|_2 + \left\|\boldsymbol{v}^t - \boldsymbol{v}^{t-1}\right\|_2 . \tag{17}$$

(In the SCI task) Consider the proposed formulation

$$\boldsymbol{x}^t = \arg\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|_2^2 + \gamma \left\|\boldsymbol{x} - \mathcal{D}\left(\boldsymbol{x}^{t-1}, \sigma\right)\right\|_2^2, \quad (18)$$

based on the first order optimality of Eqn. (18), we have

$$\nabla f(\boldsymbol{x}) + \gamma\left(\boldsymbol{x} - \mathcal{D}\left(\boldsymbol{x}^{t-1}, \sigma\right)\right) = 0. \quad (19)$$

We assume that the optimal solution is $\boldsymbol{x}^t$ at the $t-$th iteration. Substituting this solution into the above equation yields

$$\boldsymbol{x}^t - \mathcal{D}\left(\boldsymbol{x}^{t-1}, \sigma\right) = -\frac{\nabla f(\boldsymbol{x}^t)}{\gamma^{t-1}}. \quad (20)$$

At the $t-$th iteration, the value of $\gamma$ for updating image is $\gamma^{t-1}$. Using the **Lemma 2**, we have

$$\left\|\boldsymbol{x}^t - \mathcal{D}\left(\boldsymbol{x}^{t-1}, \sigma\right)\right\|_2 \leq \frac{\sqrt{\lambda_2}\left(2\sqrt{\lambda_1 N} + \epsilon\right)}{\gamma^{t-1}}. \quad (21)$$

Based on this bound and the bounded property of BSTNet, we obtain

$$\begin{aligned} &\left\|\boldsymbol{x}^t - \boldsymbol{x}^{t-1}\right\|_2 \\ =&\left\|\boldsymbol{x}^t - \mathcal{D}\left(\boldsymbol{x}^{t-1}, \sigma\right) + \mathcal{D}\left(\boldsymbol{x}^{t-1}, \sigma\right) - \boldsymbol{x}^{t-1}\right\|_2 \\ \leq&\left\|\boldsymbol{x}^{(t)} - \mathcal{D}\left(\boldsymbol{x}^{t-1}, \sigma\right)||_2 + ||\mathcal{D}\left(\boldsymbol{x}^{t-1}, \sigma\right) - \boldsymbol{x}^{t-1}\right\|_2 \\ \leq&\frac{\sqrt{\lambda_2}\left(2\sqrt{\lambda_1 N} + \epsilon\right)}{\gamma^{t-1}} + C\sigma^{t-1}. \end{aligned} \quad (22)$$

Based on this, we have

$$\begin{aligned} &\left|\left|\boldsymbol{v}^t - \boldsymbol{v}^{t-1}\right|\right|_2 \\ =&\left|\left|\mathcal{D}_\sigma\left(\boldsymbol{x}^t\right) - \mathcal{D}(\boldsymbol{x}^{t-1}, \sigma)\right|\right|_2 \\ =&||\mathcal{D}_\sigma\left(\boldsymbol{x}^t\right) - \boldsymbol{x}^t \\ &+ \boldsymbol{x}^t - \boldsymbol{x}^{t-1} + \boldsymbol{x}^{t-1} - \mathcal{D}(\boldsymbol{x}^{t-1}, \sigma)||_2 \\ \leq&\left|\left|\mathcal{D}_\sigma\left(\boldsymbol{x}^t\right) - \boldsymbol{x}^t\right|\right|_2 \\ &+ \left|\left|\boldsymbol{x}^t - \boldsymbol{x}^{t-1}\right|\right|_2 + \left|\left|\boldsymbol{x}^{t-1} - \mathcal{D}(\boldsymbol{x}^{t-1}, \sigma)\right|\right|_2. \end{aligned} \quad (23)$$

The inequality holds true due to the triangle inequality. Based on the bounded property of BSTNet and (22), we have

$$\begin{aligned} &\left|\left|\boldsymbol{v}^t - \boldsymbol{v}^{t-1}\right|\right|_2 \\ \leq&\left|\left|\mathcal{D}_\sigma\left(\boldsymbol{x}^t\right) - \boldsymbol{x}^t\right|\right|_2 + \left|\left|\boldsymbol{x}^t - \boldsymbol{x}^{t-1}\right|\right|_2 \\ &+ \left|\left|\boldsymbol{x}^{t-1} - \mathcal{D}\left(\boldsymbol{x}^{t-1}, \sigma\right)\right|\right|_2 \\ \leq&C\sigma^t + \frac{\sqrt{\lambda_2}\left(2\sqrt{\lambda_1 N} + \epsilon\right)}{\gamma^{t-1}} + C\sigma^{t-1} + C\sigma^{t-1} \\ \leq&3C\sigma^{t-1} + \frac{\sqrt{\lambda_2}\left(2\sqrt{\lambda_1 N} + \epsilon\right)}{\gamma^{t-1}}. \end{aligned} \quad (24)$$

Using the distance function defined in Eqn. (17), it is easily to see from Eqn. (22) and Eqn. (24) that

$$\begin{aligned} D\left(\boldsymbol{\theta}^t, \boldsymbol{\theta}^{t-1}\right) \leq& \frac{\sqrt{\lambda_2}\left(2\sqrt{\lambda_1 N} + \epsilon\right)}{\gamma^{t-1}} + C\sigma^{t-1} \\ &+ 3C\sigma^{t-1} + \frac{\sqrt{\lambda_2}\left(2\sqrt{\lambda_1 N} + \epsilon\right)}{\gamma^{t-1}} \\ =& \frac{\sqrt{\lambda_2}\left(4\sqrt{\lambda_1 N} + 2\epsilon\right)}{\gamma^{t-1}} + 4C\sigma^{t-1}, \end{aligned} \quad (25)$$

Since $\frac{1}{1+\gamma^t} < 1$ and $\sigma$ diminishes grandually, when $t \to \infty$, $\gamma^t \to \infty$ and $\sigma \to 0$, then $\left\|\boldsymbol{x}^t - \boldsymbol{x}^{t-1}\right\|_2 \to 0$, $\left\|\boldsymbol{v}^t - \boldsymbol{v}^{t-1}\right\|_2 \to 0$, and $D\left(\boldsymbol{\theta}^t, \boldsymbol{\theta}^{t-1}\right) \to 0$.

(In the CSMRI task) Consider the proposed formulation

$$\boldsymbol{x}^t = \arg\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{F}_u\boldsymbol{x}\|_2^2 + \gamma \left\|\boldsymbol{x} - \mathcal{D}\left(\boldsymbol{x}^{t-1}, \sigma\right)\right\|_2^2, \quad (26)$$

based on the first order optimality of Eqn. (26), we have

$$\nabla f(\boldsymbol{x}) + \gamma\left(\boldsymbol{x} - \mathcal{D}\left(\boldsymbol{x}^{t-1}, \sigma\right)\right) = 0. \quad (27)$$

We assume that the optimal solution is $\boldsymbol{x}^t$ at the $t-$th iteration. Substituting this solution into the above equation yields

$$\boldsymbol{x}^t - \mathcal{D}\left(\boldsymbol{x}^{t-1}, \sigma\right) = -\frac{\nabla f(\boldsymbol{x}^t)}{\gamma^{t-1}}. \quad (28)$$

At the $t-$th iteration, the value of $\gamma$ for updating image is $\gamma^{t-1}$. Using the **Lemma 2**, we have

$$\left\|\boldsymbol{x}^t - \mathcal{D}\left(\boldsymbol{x}^{t-1}, \sigma\right)\right\|_2 \leq \frac{\left(2\sqrt{N} + \epsilon\right)}{\gamma^{t-1}}. \quad (29)$$

Based on this bound and the bounded property of BSTNet, we obtain

$$\begin{aligned} &\left\|\boldsymbol{x}^t - \boldsymbol{x}^{t-1}\right\|_2 \\ =&\left\|\boldsymbol{x}^t - \mathcal{D}\left(\boldsymbol{x}^{t-1}, \sigma\right) + \mathcal{D}\left(\boldsymbol{x}^{t-1}, \sigma\right) - \boldsymbol{x}^{t-1}\right\|_2 \\ \leq&\left\|\boldsymbol{x}^{(t)} - \mathcal{D}\left(\boldsymbol{x}^{t-1}, \sigma\right)||_2 + ||\mathcal{D}\left(\boldsymbol{x}^{t-1}, \sigma\right) - \boldsymbol{x}^{t-1}\right\|_2 \\ \leq&\frac{\left(2\sqrt{N} + \epsilon\right)}{\gamma^{t-1}} + C\sigma^{t-1}. \end{aligned} \quad (30)$$

Based on this, we have

$$\begin{aligned} &\left|\left|\boldsymbol{v}^t - \boldsymbol{v}^{t-1}\right|\right|_2 \\ =&\left|\left|\mathcal{D}_\sigma\left(\boldsymbol{x}^t\right) - \mathcal{D}(\boldsymbol{x}^{t-1}, \sigma)\right|\right|_2 \\ =&||\mathcal{D}_\sigma\left(\boldsymbol{x}^t\right) - \boldsymbol{x}^t \\ &+ \boldsymbol{x}^t - \boldsymbol{x}^{t-1} + \boldsymbol{x}^{t-1} - \mathcal{D}(\boldsymbol{x}^{t-1}, \sigma)||_2 \\ \leq&\left|\left|\mathcal{D}_\sigma\left(\boldsymbol{x}^t\right) - \boldsymbol{x}^t\right|\right|_2 \\ &+ \left|\left|\boldsymbol{x}^t - \boldsymbol{x}^{t-1}\right|\right|_2 + \left|\left|\boldsymbol{x}^{t-1} - \mathcal{D}(\boldsymbol{x}^{t-1}, \sigma)\right|\right|_2. \end{aligned} \quad (31)$$

The inequality holds true due to the triangle inequality. By utilizing the bounded property of BSTNet and (30), we have

$$\begin{aligned} &\left|\left|\boldsymbol{v}^t - \boldsymbol{v}^{t-1}\right|\right|_2 \\ \leq&\left|\left|\mathcal{D}_\sigma\left(\boldsymbol{x}^t\right) - \boldsymbol{x}^t\right|\right|_2 + \left|\left|\boldsymbol{x}^t - \boldsymbol{x}^{t-1}\right|\right|_2 \\ &+ \left|\left|\boldsymbol{x}^{t-1} - \mathcal{D}\left(\boldsymbol{x}^{t-1}, \sigma\right)\right|\right|_2 \\ \leq&C\sigma^t + \frac{\left(2\sqrt{N} + \epsilon\right)}{\gamma^{t-1}} + C\sigma^{t-1} + C\sigma^{t-1} \\ \leq&3C\sigma^{t-1} + \frac{\left(2\sqrt{N} + \epsilon\right)}{\gamma^{t-1}}. \end{aligned} \quad (32)$$

Using the distance function defined in Eqn. (17), it is easily to see from Eqn. (30) and Eqn. (32) that

$$\begin{aligned} D\left(\boldsymbol{\theta}^t, \boldsymbol{\theta}^{t-1}\right) \leq& \frac{\left(2\sqrt{N} + \epsilon\right)}{\gamma^{t-1}} + C\sigma^{t-1} \\ &+ 3C\sigma^{t-1} + \frac{\left(2\sqrt{N} + \epsilon\right)}{\gamma^{t-1}} \\ =& \frac{\left(4\sqrt{N} + 2\epsilon\right)}{\gamma^{t-1}} + 4C\sigma^{t-1}, \end{aligned} \quad (33)$$

Since $\gamma^t = \eta\gamma^{t-1} = \eta^{t-1}\gamma^0$ and $\sigma$ diminishes grandually, where $\eta > 1$, when $t \to \infty$, $\gamma^t \to \infty$ and $\sigma \to 0$, then $\left\|\boldsymbol{x}^t - \boldsymbol{x}^{t-1}\right\|_2 \to 0$, $\left\|\boldsymbol{v}^t - \boldsymbol{v}^{t-1}\right\|_2 \to 0$, and $D\left(\boldsymbol{\theta}^t, \boldsymbol{\theta}^{t-1}\right) \to 0$.

This completes the proof of **Theorem 2** in different tasks.

## IV. The visualizations of 10 scenes


Fig. 1: The RGB images on 10 scenes of KAIST dataset.

## References

[1] E. Ryu, J. L. Liu, S. C. Wang, X. H. Chen, Z. Y. Wang, W. T. Yin, Plug-and-play methods provably converge with properly trained denoisers, in: Proceedings of the International Conference on Machine Learning (ICML), 2019, pp. 5546–5557.

[2] B. Shi, K. Liu, Regularization by multiple dual frames for compressed sensing magnetic resonance imaging with convergence analysis., IEEE/CAA Journal of Automatica Sinica 10 (11) (2023) 2136–2153. doi:10.1109/JAS.2023.123543.

[3] S. H. Chan, X. Wang, O. A. Elgendy, Plug-and-play ADMM for image restoration: Fixed-point convergence and applications, IEEE Transactions on Computational Imaging. 3 (1) (2017) 84–98. doi:10.1109/TCI.2016.2629286.