# Supplementary Materials of "HLRTF: Hierarchical Low-Rank Tensor Factorization for Inverse Problems in Multi-Dimensional Imaging"

Yisi Luo[1] Xile Zhao[1*] Deyu Meng[2,3] Taixiang Jiang[4]
[1]University of Electronic Science and Technology of China, Chengdu, China
[2]Xi'an Jiaotong University, Xi'an, China
[3]Peng Cheng Laboratory, Shenzhen, China
[4]Southwestern University of Finance and Economics, Chengdu, China
yisiluo1221@foxmail.com, xlzhao122003@163.com, dymeng@mail.xjtu.edu.cn, taixiangjiang@gmail.com

## Proof of Theorem 2

*Proof.* We first prove **(i)**. Assume that $\mathrm{rank}_h(\mathcal{X}) = r$. According to Definition 3, we have $\mathrm{rank}(f(\mathcal{X})^{(i)}) = r_i \le r$ for $i = 1, 2, \cdots, n_3$. Therefore, $f(\mathcal{X})^{(i)}$ can be factorized as $f(\mathcal{X})^{(i)} = \mathbf{A}_i\mathbf{B}_i$, where $\mathbf{A}_i \in \mathbb{R}^{n_1\times r_i}$, $\mathbf{B}_i \in \mathbb{R}^{r_i\times n_2}$ and they meet $\mathrm{rank}(\mathbf{A}_i) = \mathrm{rank}(\mathbf{B}_i) = r_i$. Let $\overline{\mathbf{A}}_i = [\mathbf{A}_i, \mathbf{0}] \in \mathbb{R}^{n_1\times r}$ and $\overline{\mathbf{B}}_i = [\mathbf{B}_i; \mathbf{0}] \in \mathbb{R}^{r\times n_2}$, then we have $f(\mathcal{X})^{(i)} = \overline{\mathbf{A}}_i\overline{\mathbf{B}}_i$. Let $\mathcal{A} = g(\overline{\mathcal{A}}) \in \mathbb{R}^{n_1\times r\times n_3}$ and $\mathcal{B} = g(\overline{\mathcal{B}}) \in \mathbb{R}^{r\times n_2\times n_3}$ be two tensors, where $g(\cdot)$ is the inverse DNN of $f(\cdot)$, $\overline{\mathcal{A}}^{(i)} = \overline{\mathbf{A}}_i$, and $\overline{\mathcal{B}}^{(i)} = \overline{\mathbf{B}}_i$ ($i = 1, 2, \cdots, n_3$). Then, we can have

$$\begin{aligned}\mathcal{A} *_f \mathcal{B} &= g(f(\mathcal{A})\triangle f(\mathcal{B}))\\ &= g(f(g(\overline{\mathcal{A}}))\triangle f(g(\overline{\mathcal{B}})))\\ &= g(\overline{\mathcal{A}}\triangle\overline{\mathcal{B}})\\ &= g(f(\mathcal{X}))\\ &= \mathcal{X}.\end{aligned} \tag{1}$$

Since $\mathrm{rank}_h(\mathcal{X}) = r$, there exists $j \in \{1, 2, \cdots, n_3\}$ such that $\mathrm{rank}(f(\mathcal{X})^{(j)}) = r$ holds. Thus, $\mathrm{rank}(\overline{\mathbf{A}}_j) = \mathrm{rank}(\overline{\mathbf{B}}_j) = r$ holds. According to Definition 3, we have $\mathrm{rank}_h(\mathcal{A}) = \mathrm{rank}_h(\mathcal{B}) = r$ holds.

Then, we prove the property **(ii)**. Assume that $\mathcal{Y} \in \mathbb{R}^{n_1\times n_2\times n_3}$ and $\mathcal{Z} \in \mathbb{R}^{n_2\times n_4\times n_3}$. Then,

$$\begin{aligned}&\mathrm{rank}_h(\mathcal{Y} *_f \mathcal{Z})\\ &= \max_{i=1,2,\cdots,n_3}\{\mathrm{rank}(f(\mathcal{Y} *_f \mathcal{Z})^{(i)})\}\\ &= \max_{i=1,2,\cdots,n_3}\{\mathrm{rank}((f(\mathcal{Y})\triangle f(\mathcal{Z}))^{(i)})\}\\ &= \max_{i=1,2,\cdots,n_3}\{\mathrm{rank}(f(\mathcal{Y})^{(i)}f(\mathcal{Z})^{(i)})\}\\ &\le \max_{i=1,2,\cdots,n_3}\{\min\{\mathrm{rank}(f(\mathcal{Y})^{(i)}), \mathrm{rank}(f(\mathcal{Z})^{(i)})\}\}\\ &\le \max_{i=1,2,\cdots,n_3}\{\mathrm{rank}(f(\mathcal{Y})^{(i)})\} = \mathrm{rank}_h(\mathcal{Y}).\end{aligned} \tag{2}$$

Similarly, we can have $\mathrm{rank}_h(\mathcal{Y} *_f \mathcal{Z}) \le \mathrm{rank}_h(\mathcal{Z})$. Thus, the following inequality holds: $\mathrm{rank}_h(\mathcal{Y} *_f \mathcal{Z}) \le \min\{\mathrm{rank}_h(\mathcal{Y}), \mathrm{rank}_h(\mathcal{Z})\}$. □

## Proof of Lemma 2

*Proof.* Assume that $L(\mathcal{X}, \mathcal{O}) = \|(\mathcal{X} - \mathcal{O})_\Omega\|_F^2$, where $\mathcal{X} = g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}})$. Here, we use $\mathcal{X}(i, j, k)$ or $\mathcal{X}_{ijk}$ to denote the $i, j, k$-th element of $\mathcal{X}$.

We first prove **(i)**. Suppose that $(i, b, c) \notin \Omega$ for arbitrary $b, c$. Then for arbitrary $v, w$, the gradient of $L$ on the $i, v, w$-th element of $\hat{\mathcal{A}}$ is

$$\begin{aligned}&\frac{\partial L(\mathcal{X}, \mathcal{O})}{\partial\hat{\mathcal{A}}(i, v, w)}\\ &= \sum_{a,b,c}\frac{\partial L}{\partial(g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}}))_{abc}}\frac{\partial(g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}}))_{abc}}{\partial\hat{\mathcal{A}}(i, v, w)}\\ &= \sum_{(a,b,c)\in\Omega} 2((g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}}) - \mathcal{O})_{abc})\frac{\partial(g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}}))_{abc}}{\partial\hat{\mathcal{A}}(i, v, w)}.\end{aligned} \tag{3}$$

Note that

$$\begin{aligned}(g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}}))_{abc} = \sum_{j_1=1}^{n_3}\mathbf{H}_{k_{cj_1}}\sigma(\sum_{j_2=1}^{n_3}\mathbf{H}_{k-1_{j_1j_2}}\cdots\\ \cdots\sigma(\sum_{j_k=1}^{n_3}\mathbf{H}_{1_{j_{k-1}j_k}}(\sum_{s=1}^{r}\hat{\mathcal{A}}_{asj_k}\hat{\mathcal{B}}_{sbj_k}))).\end{aligned} \tag{4}$$

Therefore, for arbitrary $a$ such that $a \neq i$, we have

$$\frac{\partial(g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}}))_{abc}}{\partial\hat{\mathcal{A}}(i, v, w)} = 0. \tag{5}$$

[*]Corresponding author

Thus,

$$\frac{\partial L(\mathcal{X},\mathcal{O})}{\partial \hat{\mathcal{A}}(i,v,w)} = \sum_{(i,b,c)\in\Omega} 2((g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}})-\mathcal{O})_{ibc}) \frac{\partial (g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}}))_{ibc}}{\partial \hat{\mathcal{A}}(i,v,w)}. \tag{6}$$

Since $(i,b,c)\notin\Omega$ for arbitrary $b,c$, we have

$$\frac{\partial L(\mathcal{X},\mathcal{O})}{\partial \hat{\mathcal{A}}(i,v,w)} = 0, v=1,2,\cdots,r,\ w=1,2,\cdots,n_3. \tag{7}$$

Then, we prove **(ii)**. Suppose that $(a,i,c)\notin\Omega$ for arbitrary $a,c$. Then for arbitrary $u,w$, the gradient of $L$ on the $u,i,w$-th element of $\hat{\mathcal{B}}$ is

$$\frac{\partial L(\mathcal{X},\mathcal{O})}{\partial \hat{\mathcal{B}}(u,i,w)} = \sum_{(a,b,c)\in\Omega} 2((g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}})-\mathcal{O})_{abc}) \frac{\partial (g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}}))_{abc}}{\partial \hat{\mathcal{B}}(u,i,w)}. \tag{8}$$

According to Eq. (4), for arbitrary $b$ such that $b\neq i$, we have

$$\frac{\partial (g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}}))_{abc}}{\partial \hat{\mathcal{B}}(u,i,w)} = 0. \tag{9}$$

Thus,

$$\frac{\partial L(\mathcal{X},\mathcal{O})}{\partial \hat{\mathcal{B}}(u,i,w)} = \sum_{(a,i,c)\in\Omega} 2((g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}})-\mathcal{O})_{aic}) \frac{\partial (g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}}))_{aic}}{\partial \hat{\mathcal{B}}(u,i,w)}. \tag{10}$$

Since $(a,i,c)\notin\Omega$ for arbitrary $a,c$, we have

$$\frac{\partial L(\mathcal{X},\mathcal{O})}{\partial \hat{\mathcal{B}}(u,i,w)} = 0, u=1,2,\cdots,r,\ w=1,2,\cdots,n_3. \tag{11}$$

Then, we prove **(iii)**. Suppose that $(a,b,i)\notin\Omega$ for arbitrary $a,b$. Then for arbitrary $v$, the gradient of $L$ on the $i,v$-th element of $\mathbf{H}_k$ is

$$\frac{\partial L(\mathcal{X},\mathcal{O})}{\partial \mathbf{H}_k(i,v)} = \sum_{(a,b,c)\in\Omega} 2((g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}})-\mathcal{O})_{abc}) \frac{\partial (g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}}))_{abc}}{\partial \mathbf{H}_k(i,v)}. \tag{12}$$

According to Eq. (4), for arbitrary $c$ such that $c\neq i$, we have

$$\frac{\partial (g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}}))_{abc}}{\partial \mathbf{H}_k(i,v)} = 0. \tag{13}$$

Thus,

$$\frac{\partial L(\mathcal{X},\mathcal{O})}{\partial \mathbf{H}_k(i,v)} = \sum_{(a,b,i)\in\Omega} 2((g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}})-\mathcal{O})_{abi}) \frac{\partial (g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}}))_{abi}}{\partial \mathbf{H}_k(i,v)}. \tag{14}$$

Since $(a,b,i)\notin\Omega$ for arbitrary $a,b$, we have

$$\frac{\partial L(\mathcal{X},\mathcal{O})}{\partial \mathbf{H}_k(i,v)} = 0, v=1,2,\cdots,n_3. \tag{15}$$

□

## Proof of Theorem 3

*Proof.* Suppose that $\mathcal{X} = g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}})$. Since $\sigma^{-1}(\cdot)$ is the inverse function of the LeakyReLU function [3], we have $\sigma^{-1}(0)=0$. Since $\sigma^{-1}(\cdot)$ is Lipschitz continuous, there exists $P\geq 0$ such that $\|\sigma^{-1}(\mathcal{A})\|_{\ell_1} = \|\sigma^{-1}(\mathcal{A}) - \sigma^{-1}(\mathbf{0})\|_{\ell_1} \leq P\|\mathcal{A}-\mathbf{0}\|_{\ell_1} = P\|\mathcal{A}\|_{\ell_1}$ holds for arbitrary $\mathcal{A}$. Then

$$\begin{aligned}
&\|\nabla_x \mathcal{X}\|_{\ell_1} \\
=&\|\mathcal{X}_{(1:n_1-1,:,:)} - \mathcal{X}_{(2:n_1,:,:)}\|_{\ell_1} \\
=&\|\sigma^{-1}(\cdots\sigma^{-1}(\sigma^{-1}((\hat{\mathcal{A}}_{(1:n_1-1,:,:)}\triangle\hat{\mathcal{B}})\times_3\mathbf{H}_1)\times_3 \\
&\mathbf{H}_2)\cdots\times_3\mathbf{H}_{k-1})\times_3\mathbf{H}_k - \sigma^{-1}(\cdots\sigma^{-1}(\sigma^{-1}(( \\
&\hat{\mathcal{A}}_{(2:n_1,:,:)}\triangle\hat{\mathcal{B}})\times_3\mathbf{H}_1)\times_3\mathbf{H}_2)\cdots\times_3\mathbf{H}_{k-1})\times_3\mathbf{H}_k\|_{\ell_1} \\
\leq&\|\sigma^{-1}(\cdots\sigma^{-1}(\sigma^{-1}((\hat{\mathcal{A}}_{(1:n_1-1,:,:)}\triangle\hat{\mathcal{B}})\times_3\mathbf{H}_1)\times_3 \\
&\mathbf{H}_2)\cdots\times_3\mathbf{H}_{k-1}) - \sigma^{-1}(\cdots\sigma^{-1}(\sigma^{-1}((\hat{\mathcal{A}}_{(2:n_1,:,:)}\triangle \\
&\hat{\mathcal{B}})\times_3\mathbf{H}_1)\times_3\mathbf{H}_2)\cdots\times_3\mathbf{H}_{k-1})\|_{\ell_1}\|\mathbf{H}_k\|_{\ell_1} \\
\leq&P\|\sigma^{-1}(\cdots\sigma^{-1}(\sigma^{-1}((\hat{\mathcal{A}}_{(1:n_1-1,:,:)}\triangle\hat{\mathcal{B}})\times_3\mathbf{H}_1)\times_3 \\
&\mathbf{H}_2)\cdots\times_3\mathbf{H}_{k-2})\times_3\mathbf{H}_{k-1} - \sigma^{-1}(\cdots\sigma^{-1}(\sigma^{-1}(( \\
&\hat{\mathcal{A}}_{(2:n_1,:,:)}\triangle\hat{\mathcal{B}})\times_3\mathbf{H}_1)\times_3\mathbf{H}_2)\cdots\times_3\mathbf{H}_{k-2})\times_3 \\
&\mathbf{H}_{k-1}\|_{\ell_1}\|\mathbf{H}_k\|_{\ell_1} \\
\leq&P^{k-1}\|\mathbf{H}_1\|_{\ell_1}\|\mathbf{H}_2\|_{\ell_1}\cdots\|\mathbf{H}_k\|_{\ell_1}\|(\hat{\mathcal{A}}_{(1:n_1-1,:,:)}\triangle\hat{\mathcal{B}}) - \\
&(\hat{\mathcal{A}}_{(2:n_1,:,:)}\triangle\hat{\mathcal{B}})\|_{\ell_1} \\
\leq&P^{k-1}\|\mathbf{H}_1\|_{\ell_1}\|\mathbf{H}_2\|_{\ell_1}\cdots\|\mathbf{H}_k\|_{\ell_1}\|\hat{\mathcal{B}}\|_{\ell_1}\|\hat{\mathcal{A}}_{(1:n_1-1,:,:)} - \\
&\hat{\mathcal{A}}_{(2:n_1,:,:)}\|_{\ell_1} \\
=&P^{k-1}\|\mathbf{H}_1\|_{\ell_1}\|\mathbf{H}_2\|_{\ell_1}\cdots\|\mathbf{H}_k\|_{\ell_1}\|\hat{\mathcal{B}}\|_{\ell_1}\|\nabla_x\hat{\mathcal{A}}\|_{\ell_1}.
\end{aligned} \tag{16}$$

Since $\{\mathbf{H}_j\}_{j=1}^k$ and $\hat{\mathcal{B}}$ are bounded, we have

$$\|\nabla_x\mathcal{X}\|_{\ell_1} \leq J_1\|\nabla_x\hat{\mathcal{A}}\|_{\ell_1}, \tag{17}$$

where $J_1 = P^{k-1}\|\mathbf{H}_1\|_{\ell_1}\|\mathbf{H}_2\|_{\ell_1}\cdots\|\mathbf{H}_k\|_{\ell_1}\|\hat{\mathcal{B}}\|_{\ell_1}$ is a constant.

Figure 1. Tensor completion results vs. $\gamma$ and $r$ (*Pavia*, SR=0.1).

Similar to (16), we have

$$\begin{aligned}\|\nabla_y \mathcal{X}\|_{\ell_1} \leq & P^{k-1}\|\mathbf{H}_1\|_{\ell_1}\|\mathbf{H}_2\|_{\ell_1}\cdots\|\mathbf{H}_k\|_{\ell_1}\|\hat{\mathcal{A}}\|_{\ell_1}\|\nabla_y \hat{\mathcal{B}}\|_{\ell_1} \\ = & J_2\|\nabla_y \hat{\mathcal{B}}\|_{\ell_1},\end{aligned} \tag{18}$$

where $J_2 = P^{k-1}\|\mathbf{H}_1\|_{\ell_1}\|\mathbf{H}_2\|_{\ell_1}\cdots\|\mathbf{H}_k\|_{\ell_1}\|\hat{\mathcal{A}}\|_{\ell_1}$ is a constant.

Next, we prove the third inequality. Since

$$\begin{aligned}&\|\nabla_z \mathcal{X}\|_{\ell_1} \\ =&\|\mathcal{X}_{(:,:,1:n_3-1)} - \mathcal{X}_{(:,:,2:n_3)}\|_{\ell_1} \\ =&\|\sigma^{-1}(\cdots\sigma^{-1}(\sigma^{-1}((\hat{\mathcal{A}}\triangle\hat{\mathcal{B}})\times_3 \mathbf{H}_1)\times_3 \mathbf{H}_2)\cdots\times_3 \\ &\mathbf{H}_{k-1})\times_3 \mathbf{H}_{k(1:n_3-1,:)} - \sigma^{-1}(\cdots\sigma^{-1}(\sigma^{-1}((\hat{\mathcal{A}}\triangle \\ &\hat{\mathcal{B}})\times_3 \mathbf{H}_1)\times_3 \mathbf{H}_2)\cdots\times_3 \mathbf{H}_{k-1})\times_3 \mathbf{H}_{k(2:n_3,:)}\|_{\ell_1} \\ \leq&\|\sigma^{-1}(\cdots\sigma^{-1}(\sigma^{-1}((\hat{\mathcal{A}}\triangle\hat{\mathcal{B}})\times_3 \mathbf{H}_1)\times_3 \mathbf{H}_2)\cdots\times_3 \\ &\mathbf{H}_{k-1})\|_{\ell_1}\|\mathbf{H}_{k(1:n_3-1,:)} - \mathbf{H}_{k(2:n_3,:)}\|_{\ell_1} \\ \leq&P^{k-1}\|\hat{\mathcal{A}}\|_{\ell_1}\|\hat{\mathcal{B}}\|_{\ell_1}\|\mathbf{H}_1\|_{\ell_1}\|\mathbf{H}_2\|_{\ell_1}\cdots\|\mathbf{H}_{k-1}\|_{\ell_1}\|\nabla_x \mathbf{H}_k\|_{\ell_1},\end{aligned} \tag{19}$$

we have

$$\|\nabla_z \mathcal{X}\|_{\ell_1} \leq J_3\|\nabla_x \mathbf{H}_k\|_{\ell_1}, \tag{20}$$

where $J_3 = P^{k-1}\|\hat{\mathcal{A}}\|_{\ell_1}\|\hat{\mathcal{B}}\|_{\ell_1}\|\mathbf{H}_1\|_{\ell_1}\|\mathbf{H}_2\|_{\ell_1}\cdots\|\mathbf{H}_{k-1}\|_{\ell_1}$ is a constant. The proof is completed. □

## Implementation Details

In this work, all experiments are conducted on the PyTorch and MATLAB 2019a platform with an i5-9400f CPU, an RTX 3060 GPU, and 16GB RAM. As for the proposed method, we set the number of network layers $k$ as 2 in the experiments. The rank $r$ and the trade-off parameter $\gamma$ are tuned based on the highest PSNR value. The hyperparameters of compared methods are tuned to obtain the highest PSNR value. For CNN-based methods, we use the pretrained model provided by the authors.

In this work, we consider the simple fully connected network. In future work, we can consider some advanced DNNs (e.g., attention) in our framework.

The limitation of our method lies in the manual selection of the hyperparameters, i.e., the rank $r$ and the trade-off parameter $\gamma$. In experiments, we select these parameters to obtain the best PSNR value. Numerical tests in Fig. 1 show that our method is relatively insensitive to the choice of $r$ and $\gamma$. Thus, it is not difficult to choose these hyperparameters in experiments.

## More Experimental Results

Please see Figs. 3-7 for more visual results. Our method shows advantageous performance against compared state-of-the-art methods for different tasks. Moreover, in Fig. 2, we show the visualizations of the learned matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ for HSIs *WDC mall* and *Pavia* with SR=0.1. Compared with the fixed DFT matrix, our learned nonlinear transform is more flexible to represent different data.

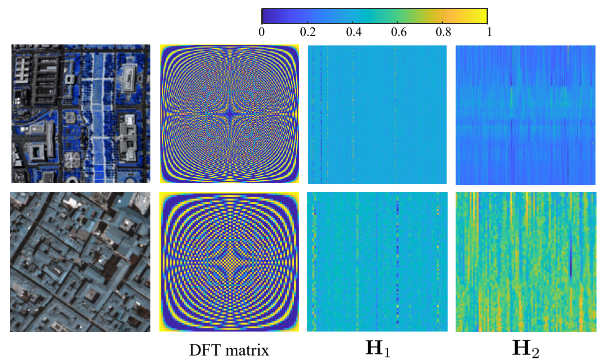

Figure 2. Given different data (e.g., the HSI *WDC mall* and *Pavia*), traditional t-SVD methods [6,21] use fixed DFT matrix for multi-dimensional image recovery. In constrast, our DeepLRTF flexibly learns different nonlinear transforms (see the learned matrices $\mathbf{H}_1$ and $\mathbf{H}_2$) for different data to help obtain better performance.

## References

[1] Chunhong Cao, Jie Yu, Chengyao Zhou, Kai Hu, Fen Xiao, and Xieping Gao. Hyperspectral image denoising via subspace-based nonlocal low-rank and sparse factorization. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(3):973–988, 2019. 4, 5

[2] Jicong Fan and Jieyu Cheng. Matrix completion by deep matrix factorization. *Neural Networks*, 98:34–41, 2018. 4, 6

[3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *ICCV, 2015*, pages 1026–1034, 2015. 2

[4] Tai-Xiang Jiang, Michael K. Ng, Xi-Le Zhao, and Ting-Zhu Huang. Framelet representation of tensor nuclear norm for third-order tensor completion. *IEEE Transactions on Image Processing*, 29:7233–7244, 2020. 4, 6

[5] Yang Liu, Xin Yuan, Jinli Suo, David J. Brady, and Qionghai Dai. Rank minimization for snapshot compressive imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(12):2990–3006, 2019. 5

[6] Canyi Lu, Jiashi Feng, Yudong Chen, Wei Liu, Zhouchen Lin, and Shuicheng Yan. Tensor robust principal component analysis with a new tensor nuclear norm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(4):925–938, 2020. 3

[7] Canyi Lu, Xi Peng, and Yunchao Wei. Low-rank tensor completion with a new tensor nuclear norm induced by invertible

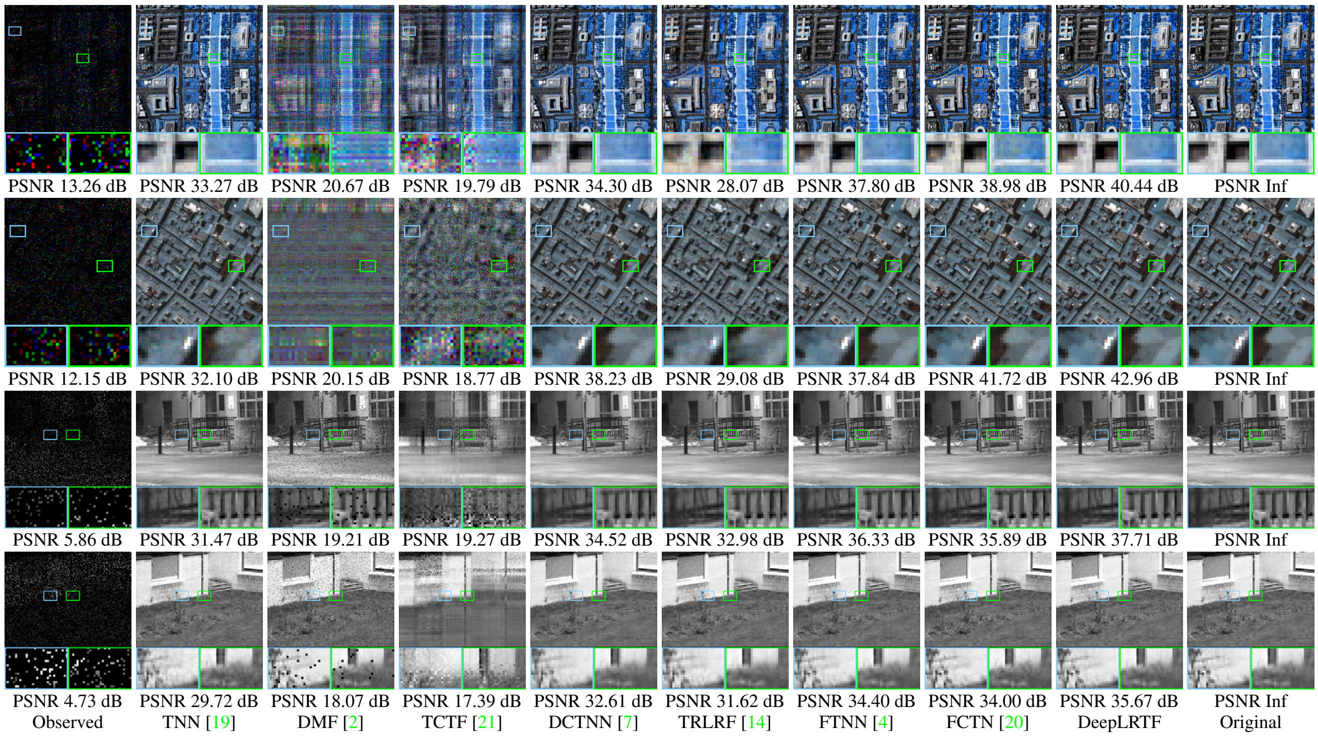


Figure 3. The results of multi-dimensional image completion by different methods on HSI *WDC mall*, HSI *Pavia*, video *Backdoor*, and video *Yard* with SR=0.1.

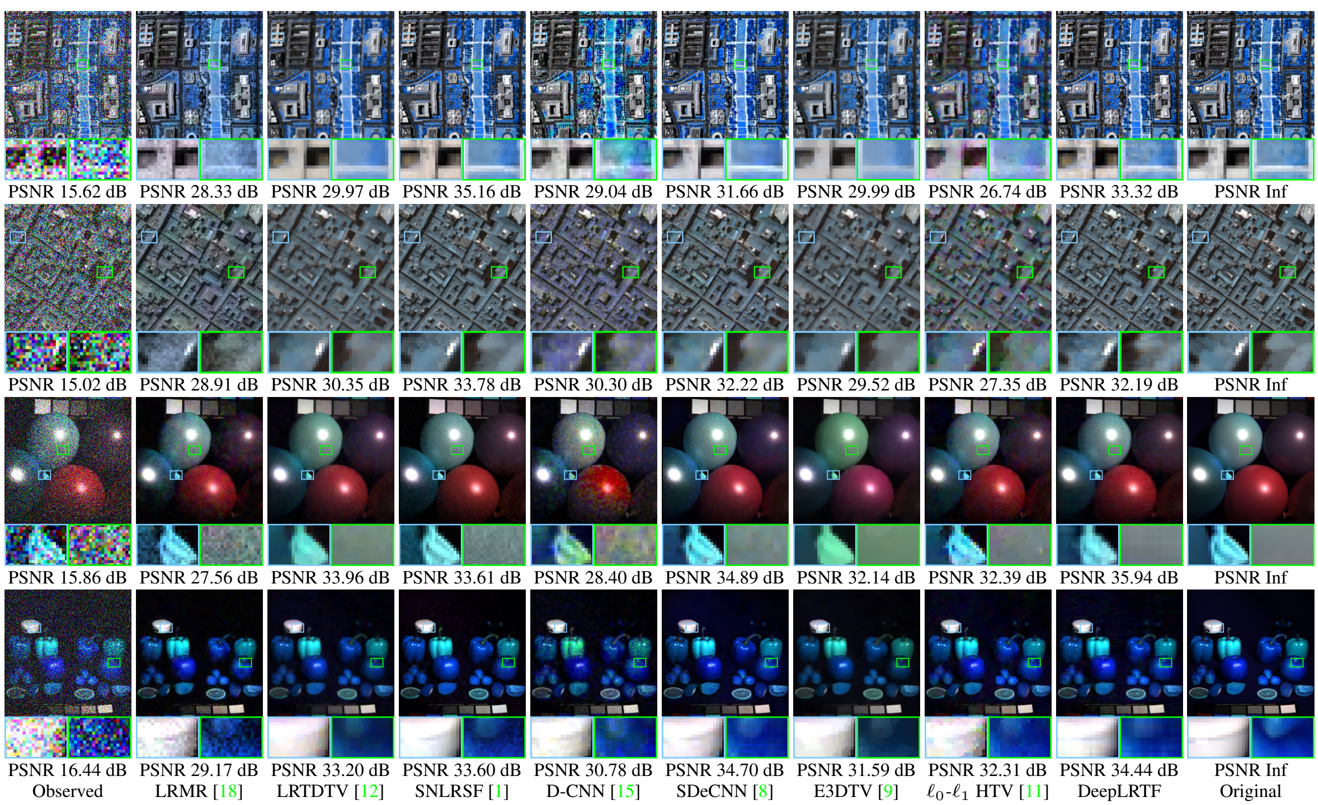


Figure 4. The results of MSI denoising by different methods on HSI *WDC mall*, HSI *Pavia*, MSI *Balloons*, and MSI *Fruits* for **Case 1**.

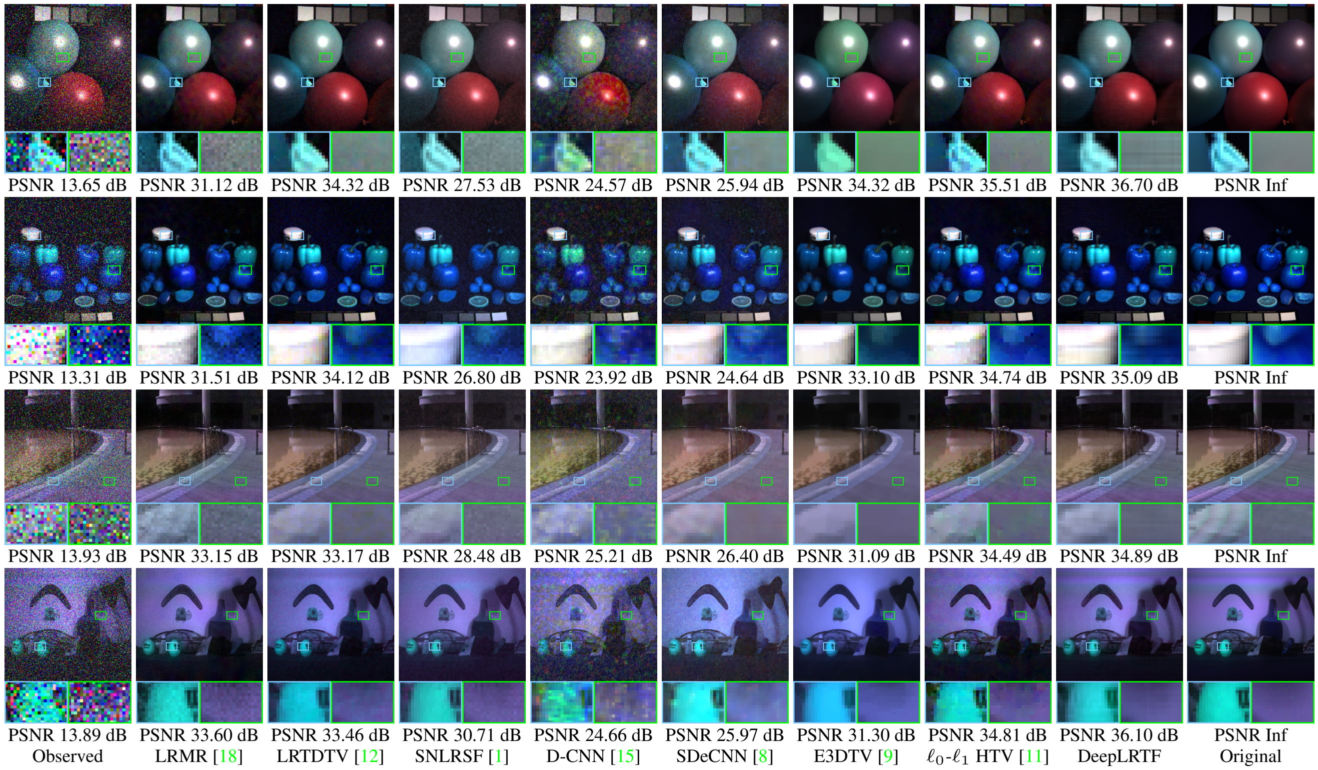


Figure 5. The results of MSI denoising by different methods on MSIs *Balloons*, *Fruits*, *Pool*, and *Doll* for **Case 2**.

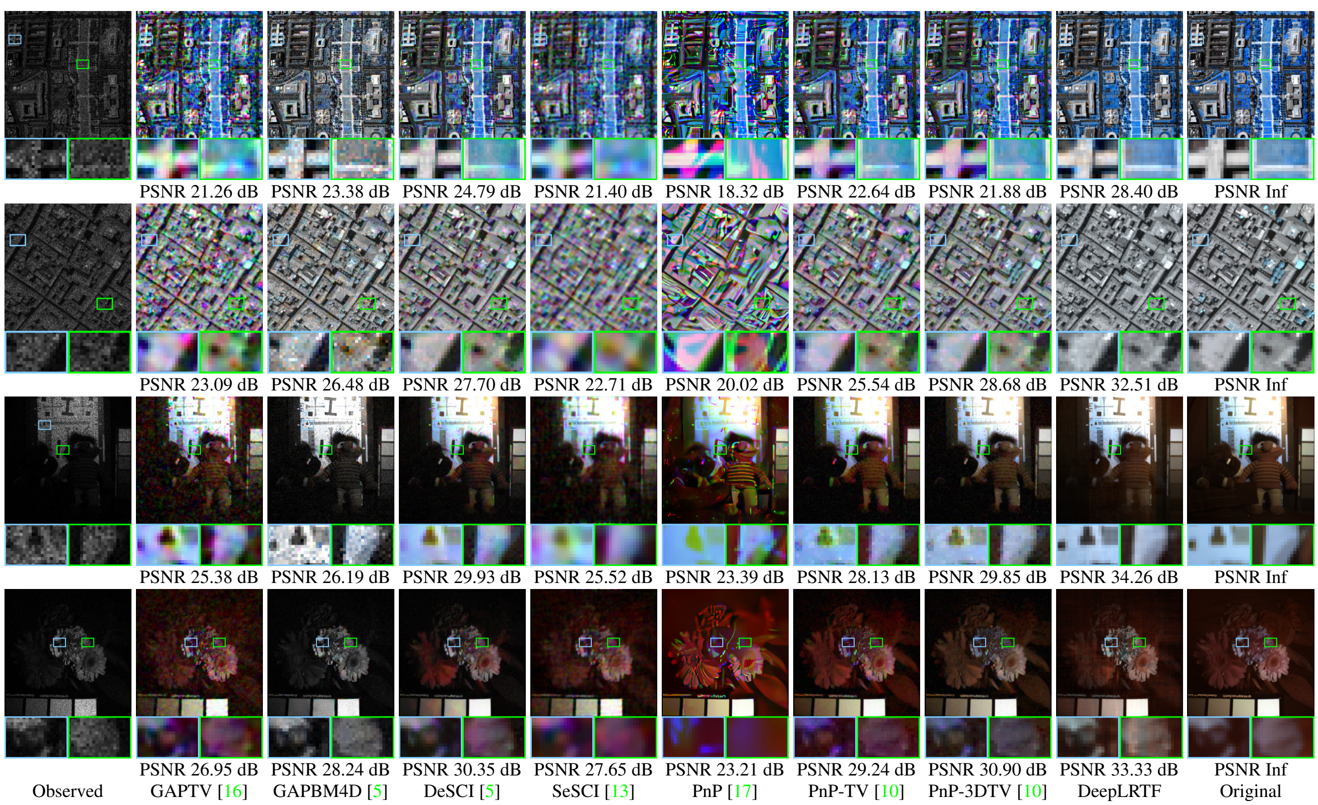


Figure 6. The results of snapshot spectral imaging by different methods on HSI *WDC mall* (SR = 0.5), HSI *Pavia* (SR = 0.5), MSI *Toy* (SR = 0.5), and MSI *Flowers* (SR = 0.3).

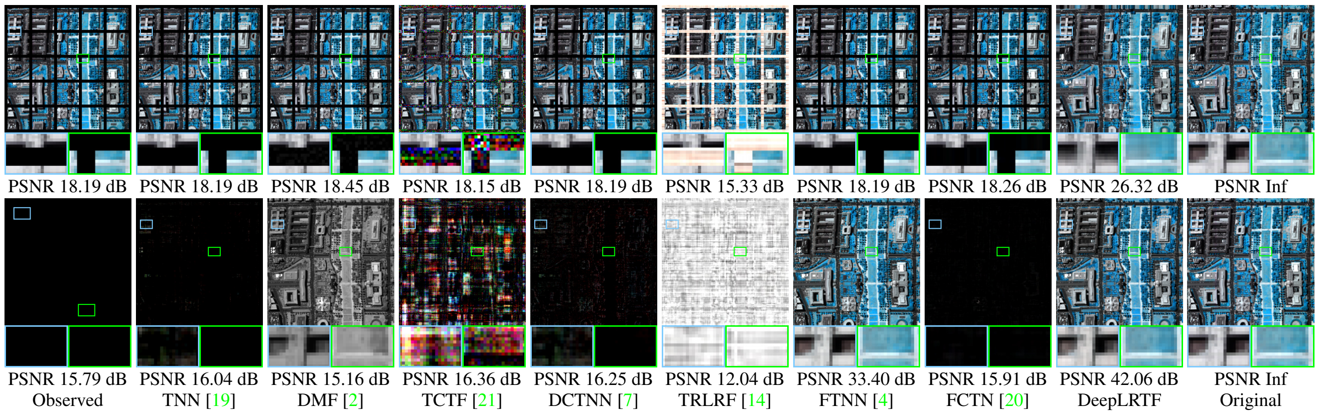


Figure 7. The results of multi-dimensional image completion by different methods on *WDC mall* with horizontal/lateral slice missing and frontal slice missing.

linear transforms. In *CVPR, 2019*, pages 5989–5997, 2019. 4, 6

[8] Alessandro Maffei, Juan M. Haut, Mercedes Eugenia Paoletti, Javier Plaza, Lorenzo Bruzzone, and Antonio Plaza. A single model CNN for hyperspectral image denoising. *IEEE Transactions on Geoscience and Remote Sensing*, 58(4):2516–2529, 2020. 4, 5

[9] Jiangjun Peng, Qi Xie, Qian Zhao, Yao Wang, Leung Yee, and Deyu Meng. Enhanced 3DTV regularization and its applications on HSI denoising and compressed sensing. *IEEE Transactions on Image Processing*, 29:7889–7903, 2020. 4, 5

[10] Haiquan Qiu, Yao Wang, and Deyu Meng. Effective snapshot compressive-spectral imaging via deep denoising and total variation priors. In *CVPR, 2021*, pages 9127–9136, 2021. 5

[11] Minghua Wang, Qiang Wang, Jocelyn Chanussot, and Danfeng Hong. $\ell_0$-$\ell_1$ hybrid total variation regularization and its applications on hyperspectral image mixed noise removal and compressed sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 2021. 4, 5

[12] Yao Wang, Jiangjun Peng, Qian Zhao, Yee Leung, Xi-Le Zhao, and Deyu Meng. Hyperspectral image restoration via total variation regularized low-rank tensor decomposition. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(4):1227–1243, 2018. 4, 5

[13] Peihao Yang, Linghe Kong, Xiao-Yang Liu, Xin Yuan, and Guihai Chen. Shearlet enhanced snapshot compressive imaging. *IEEE Transactions on Image Processing*, 29:6466–6481, 2020. 5

[14] Longhao Yuan, Chao Li, Danilo P. Mandic, Jianting Cao, and Qibin Zhao. Tensor ring decomposition with rank minimization on latent space: An efficient approach for tensor completion. In *AAAI, 2019*, pages 9151–9158, 2019. 4, 6

[15] Qiangqiang Yuan, Qiang Zhang, Jie Li, Huanfeng Shen, and Liangpei Zhang. Hyperspectral image denoising employing a spatial–spectral deep residual convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2):1205–1218, 2019. 4, 5

[16] Xin Yuan. Generalized alternating projection based total variation minimization for compressive sensing. In *ICIP, 2016*, pages 2539–2543, 2016. 5

[17] Xin Yuan, Yang Liu, Jinli Suo, and Qionghai Dai. Plug-and-play algorithms for large-scale snapshot compressive imaging. In *CVPR, 2020*, pages 1447 – 1457, 2020. 5

[18] Hongyan Zhang, Wei He, Liangpei Zhang, Huanfeng Shen, and Qiangqiang Yuan. Hyperspectral image restoration using low-rank matrix recovery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(8):4729–4743, 2014. 4, 5

[19] Zemin Zhang, Gregory Ely, Shuchin Aeron, Ning Hao, and Misha Kilmer. Novel methods for multilinear data completion and de-noising based on tensor-svd. In *CVPR, 2014*, pages 3842–3849, 2014. 4, 6

[20] Yu-Bang Zheng, Ting-Zhu Huang, Xi-Le Zhao, Qibin Zhao, and Tai-Xiang Jiang. Fully-connected tensor network decomposition and its application to higher-order tensor completion. In *AAAI, 2021*, 2021. 4, 6

[21] Pan Zhou, Canyi Lu, Zhouchen Lin, and Chao Zhang. Tensor factorization for low-rank tensor completion. *IEEE Transactions on Image Processing*, 27(3):1152–1163, 2018. 3, 4, 6