In-Orbit Lunar Satellite Image Super Resolution for Selective Data Transmission

Atal Tewari†,*+, Chennuri Prateek†,*+, Nitin Khanna†,+‡
†Electrical Engineering, Indian Institute of Technology Gandhinagar, India
‡Electrical Engineering and Computer Science, Indian Institute of Technology Bhilai, India

Abstract—Rapid technological advancements have tremendously increased the data acquisition capabilities of remote sensing satellites. However, the data utilization efficiency in satellite missions is very low. This growing data also escalates the cost required for data downlink transmission and post-processing. Selective data transmission based on in-orbit inferences will address these issues to a great extent. Therefore, to decrease the cost of the satellite mission, we propose a novel system design for selective data transmission, based on in-orbit inferences. As the resolution of images plays a critical role in making precise inferences, we also include in-orbit super-resolution (SR) in the system design. We introduce a new image reconstruction technique and a unique loss function to enable the execution of the SR model on low-power devices suitable for satellite environments. We present a residual dense non-local attention network (RDNLNA) that provides enhanced super-resolution outputs to improve the SR performance. SR experiments on Kaguya digital ortho maps (DOMs) demonstrate that the proposed SR algorithm outperforms the residual dense network (RDN) in terms of PSNR and block-sensitive PSNR by a margin of +0.1 dB and +0.19 dB, respectively. The proposed SR system consumes 48% less memory and 67% less peak instantaneous power than the standard SR model, RDN, making it more suitable for execution on a low-powered device platform.

I. INTRODUCTION

The extraction of high resolution (HR) imagery from one of its low resolution (LR) counterparts is a topic of great significance in the field of computer vision. This task, commonly referred to as single image super-resolution (SISR), is an ill-posed problem since there are always multiple HR images corresponding to a single LR image [1]. SISR has a wide range of computer vision and image processing applications such as medical diagnostics, satellite imaging, and other high-level vision tasks [1]. Specifically, in lunar satellite imaging, image resolution plays a critical role in studies such as topological and geomorphological analysis, elemental/chemical composition, and mineralogy to understand the moon’s origin and evolution.

For all the satellite subsystems, the sun acts as the primary energy source, whose energy is transferred from electromagnetic to electric using solar panels mounted around the satellite. Due to solar energy’s renewability, there is enough energy available for the satellite mission. However, there is a hard limitation on the power available due to the hardware constraints of the solar panels. Therefore, low powered devices are preferred for satellite subsystems to utilize power more effectively.

Continuous improvements in satellite imaging technology in terms of spatial and spectral resolution lead to an exponential increase in satellite data. The satellite’s strong data acquisition capacity brings tremendous pressure on the satellite data downlink transmission and post-processing [2]. Moreover, there is a vast gap between the amount of data received and the amount in actual use, implying extremely low data utilization efficiency [2]. A viable solution for such difficulties is to design a selective data transmission technique based on in-orbit satellite inferences. As image resolution plays a crucial role in making accurate inferences and high-resolution sensors are generally expensive, there is a necessity to include a super-resolution algorithm (executable on low-powered devices) in the selective data transmission system that can extract high-resolution (HR) images from the satellite’s low-resolution sensors. Therefore, there is a need for an in-orbit satellite super-resolution algorithm executable on low-powered devices suitable for satellite environments. With the rapid advancement of deep learning (DL) technology in recent years, deep learning-based SR models have been actively explored, among which single image super-resolution (SISR) has become the mainstream. Lukas et al. [3] demonstrated one of the pioneer works, which signified that deep-learning (DL) was an effective solution for remote sensing image SR. Starting from LGCNet [4] to the latest CNN-based SISR architectures like MHAN [5], SMSR [6], SR networks for remote sensing have evolved dramatically in recent times. However, most remote sensing SISR methods consume significantly high peak instantaneous power thus cannot be utilized as a part of the in-orbit SR algorithm. Though few SR DL models [7], [8], [9], [10], can be executed on low-powered devices, they experience a significant trade-off in the output quality. Therefore, utilizing such algorithms would lead to inaccurate in-orbit inferences, thus, decreasing the overall performance of the selective data transmission technique.

To address the drawbacks mentioned in the previous paragraphs, we propose a novel system design for selective data
transmission that includes an in-orbit SR step to make accurate in-orbit inferences. For the SR algorithm to be executable on low-powered devices, we follow a simple yet effective SR evaluation procedure as shown in Figure 2. We introduce a unique overlap-reconstruction mechanism that extracts patch regions with more contextual information for efficient SR image reconstruction using output patches. Along the lines of the overlap reconstruction approach, we introduce a novel and efficient loss function, mask-PSNR, that allows the model to produce enhanced results by laying higher emphasis on the regions of interest.

To improve the SR performance, we modify the residual dense network [11] into a residual dense non-local attention network that extracts the original LR image’s hierarchical features by utilizing the residual, dense, and attention mechanism to capture long-range dependencies for enhanced reconstruction. Our RDNLA is built on our proposed residual dense non-local attention block (RDNLB) shown in Figure 3. The novel and complex combination of residual, dense and non-local attention in RDNLB strengthen the model’s long-range dependent feature propagation by re-utilizing the features in a feed-forward fashion [12] which further enhances the feature representation ability, thus, improving the SR performance as a whole. A more detailed description of the proposed model can be found in Section III-B.

In summary, our main contributions are as follows:

- We identify an unsolved problem of low data utilization efficiency in satellite missions. We resolve the issue by introducing a novel system design for the selective data transmission based on in-orbit inferences.
- We propose an overlap reconstruction mechanism that extracts regions with adequate contextual information for the in-orbit super-resolution algorithm to be executable on low-powered devices suitable for satellite environments. Along the lines of overlap reconstruction, we present a new loss function, mask PSNR, that lets the SR DL model emphasize regions having rich contextual information.
- To improve the SR performance, we introduce a residual dense non-local attention network that utilizes an intricate combination of residual, dense and non-local attention blocks to effectively extract long-range dependent hierarchical features and provide enhanced SR output.
- To the best of our knowledge, this is the first work to present an SR algorithm lunar digital ortho maps (DOMs). Moreover, this is the first deep learning-based algorithm proposed for SR of lunar satellite data (Ex: DOMs, DEMs).

II. RELATED WORK

In recent years, with the increase in demand for HR images in remote sensing and the deep learning-based methods broad applicability, many researchers use deep learning-based methods for SR of remote sensing images. Lukas et al. [3] used convolution layers for SR of multi-spectral images. To learn a multi-level representation including local and global details, a local-global combined network (LGCNet) was proposed by Lei et al. [4]. A multi-perception attention network (MPSR) [13] containing enhanced residual block (ERB) and residual channel attention group (RCAG) was developed to efficiently capture the prior information and adaptively focus on informative features. Pan et al. [14] utilized residual, dense along with recursive blocks to improve the performance of super-resolved remote sensing images. A recent paper [5] uses a novel higher-order attention (HOA) mechanism was introduced in which consists of a feature extraction network and a feature refinement network to extract the input image high-frequency details effectively. The above mentioned SR solutions provide enhanced results; however, they are computationally complex and memory expensive, therefore, not suitable for executing on a low-powered device.

In recent years, there is significant research presented in the direction of low-powered device SR [7], [8], [9], [10]. Though these approaches handle the computation complexity and memory constraints to a great extent, the performance is significantly compromised compared to the standard DL architectures. Therefore, these procedures can not be utilized for selective data transmission since it would decrease the system’s overall effectiveness. Our work mainly focuses on proposing an SR algorithm that is executable on low-powered devices without compromising on the quality of the SR output.

III. METHODOLOGY

A. System Design for Selective Data Transmission

To address the problem of low data utilization efficiency in satellite missions, we present a new system design for selective data transmission based on in-orbit inferences as shown in Figure 1. The image acquisition is performed using low-resolution sensors since they are less expensive. The LR image is provided as an input to the in-orbit SR algorithm. For the SR algorithm to be executable on low-power devices, we follow a unique SR evaluation method as shown in Figure 2. More information about the in-orbit SR algorithm can be found in Sections III-B, III-C. The SR output generated through the in-orbit SR algorithm is utilized to make in-orbit inferences. The kind of inferences made entirely depends on
the goals of the satellite mission; some of them include crater detection, mineral detection, and others. Experiments on in-orbit inferences are beyond the scope of this paper. Based on the results obtained through in-orbit inferences, a decision is made whether the image is useful or not, depending on which the image is transmitted or discarded. By utilizing such a pipeline, only the valuable images are downlink transmitted, increasing the data utilization efficiency and decreasing the overall cost of downlink transmission.

B. Residual Dense Non-Local Attention Network

As shown in Figure 3, our network can be broadly divided into four parts: shallow feature extraction (SFE), residual dense non-local attention blocks (RDNLBs), dense feature blending (DFB), and up-scaling (US). Let us assume $I^{LR}$ and $I^{SR}$ to be the low resolution input and super resolution output, respectively. SFE comprises of two $3 \times 3$ convolution layers that extract low-level features. Output of the second Conv layer is represented as

$$H_0 = F_{SFE2}(H_{-1}) = F_{SFE2}(F_{SFE1}(I^{LR})), \quad (1)$$

where $F_{SFE1}$ and $F_{SFE2}$ represent the convolution operation of first and second convolution layers, respectively. $H_0$ is then used as an input to residual dense non-local attention blocks (RDNLBs) to extract dense long-range dependent hierarchical features. The output of the $p^{th}$ $(1 \leq p \leq P)$ RDNLB can be represented as

$$H_p = F_{RDNLB,p}(H_{p-1}) = F_{RDNLB,p}(F_{RDNLB,p-1}(\ldots (F_{RDNLB,1}(H_0))\ldots)), \quad (2)$$

where $F_{RDNLB,p}$ indicates the composite functions such as convolution, ReLU [15], dense connections [12] of the $p^{th}$ RDNLB. The hierarchical feature maps are further passed through the process of DFB that involves global feature blending (GFB) and global residual attention (GRA) mechanism. DFB blends the features of all the preceding RDNLBs and extracts global long-range dependent hierarchical features which are then used as input for up-scaling (US). DFB can be represented as,

$$H_{DFB} = F_{DFB}(H_{-1}, H_0, \ldots, H_P), \quad (3)$$

where $F_{DFB}$ denotes the composite function of GFB and GRA. Lastly, we utilize the local and global hierarchical features as input for US, followed by a $3 \times 3$ Conv layer for extracting the SR output.

**Residual Dense Non-Local Attention Block.** Our proposed RDNLB is shown in Figure 3. It consists of densely connected layers, local feature blending (LFB), and local residual attention (LRA) leading to a coupled memory (CM) mechanism.

The coupled memory mechanism is achieved by transferring the previous RDNLB output to each layer of the current RDNLB. Mathematically, the output of the $d^{th}$ Conv layer of the $p^{th}$ RDNLB is

$$H_{(p,d)} = \Gamma(W_{(p,d)}[H_{p-1}, H_{(p,1)}, \ldots, H_{(p,d-1)}]), \quad (4)$$

where $H_{p-1}$ and $H_p$ are the input and output of the $p^{th}$ RDNLB, each consisting of $G_b$ feature maps. $W_{(p,d)}$ is the weight vector of the $d^{th}$ Conv layer (the bias term is excluded for simplicity). $\Gamma$ represents the ReLU [15] activation function. $[H_{p-1}, H_{(p,1)}, \ldots, H_{(p,d-1)}]$ represents the concatenation of channel outputs produced by the $(p-1)^{th}$ RDNLB and the convolutional layers $1, \ldots, (d-1)$ in the $p^{th}$ RDNLB.

Local feature blending is the process of adaptively fusing the outputs of preceding RDNLB and the current RDNLB’s Conv layers. The concatenation operation itself is not viable since the size of feature-maps would change for every RDNLB. Moreover, as the growth rate $G$ increases, the network’s depth also increases, resulting in training difficulty. Therefore, to ease the training difficulty, similar to Memnet [16], RDN [11], we utilize $1 \times 1$ convolutional layers to control the channel output information adaptively. LFB can be formulated as

$$H_{p}^{D} = W_{1 \times 1}^{p}([H_{p-1}, H_{(p,1)}, \ldots, H_{(p,d-1)}, H_{(p,d)}]), \quad (5)$$

where $H_{p}^{D}$ represents the LFB output, $W_{1 \times 1}^{p}$ denotes the $1 \times 1$ convolution layer operation of the $p^{th}$ RDNLB.

Local residual attention is a combination of local residual learning and non-local attention mechanism. As there are several convolutional layers in a single RDNLB, local residual learning plays a critical role in enhancing the information flow throughout the network. The local residual learning output can be represented as

$$H_{LR}^{p} = H_{DFB}^{p} + H_{p-1}. \quad (6)$$

The main task of non-local attention mechanism, consisting of non-local attention [17] block (NLB) followed by sigmoid activation, is to grasp information of larger receptive field size, so that it is feasible to obtain sophisticated long-range dependent hierarchical features. The final output of the $d^{th}$ RDNLB can be represented as

$$H_p = H_{LR}^{p} \Lambda(F_{NLB}(H_{p-1})), \quad (7)$$

where $F_{NLB}$ and $\Lambda$ represents the non-local attention operation and sigmoid activation.
Dense Feature Blending. After extracting the local dense long-range dependent hierarchical features using RDNLBs, we further introduce DFB to exploit the features globally. Our DFB consists of GFB and GRA mechanism. Global feature blending brings out the global hierarchical features by fusing the states of all the RDNLBs. GFB can be represented as

\[ H_{\text{GFB}} = F_{\text{GFB}}([H_1, \ldots, H_{P-1}, H_P]), \]  

(8)

where \([H_1, \ldots, H_{P-1}, H_P]\) represents the concatenation of 1, \ldots, P RDNLB outputs. \(F_{\text{GFB}}\) indicates the composite function of 1 \times 1 and 3 \times 3 convolutional layers. The 1 \times 1 convolutional layer adaptively controls the output information and the 3 \times 3 convolution modifies and extracts the global hierarchical features which are further introduced into the GRA mechanism.

Global residual attention consists of global residual learning followed by the non-local attention mechanism. The global residual learning improves the information flow in the network while mitigating the vanishing gradient problem. The non-local attention mechanism is introduced to drive the network to extract long-range dependent hierarchical features globally. This encourages the model to extract details of larger receptive field, thus producing long-range dependent feature maps for improved image SR. The output of dense feature blending that serves as an input to the up-scaling block is

\[ H_{\text{DFB}} = F_{\text{GRA}}(H_{-1}, H_{\text{GFB}}), \]  

(9)

where \(F_{\text{GRA}}\) denotes the composite function of global residual learning, non-local attention, and 3 \times 3 convolution.

C. Overlap Reconstruction and Mask-PSNR

In a low-powered device, the input size plays a critical role in the amount of multiply and accumulate operations and memory access operations [19]. To be executable on such devices, we follow a patch-based prediction approach for extracting the SR outputs as shown in Figure 2. During image reconstruction, non-overlapping stacking of patches leads to significant blocking artifacts in the reconstructed output, as shown in Figure 4. It may be due to the lack of adequate contextual information at the boundaries of the patch leading to ineffective extraction of hierarchical features. To alleviate this problem, we propose an overlap reconstruction approach that provides 50% overlapping patches as input to the network. As shown in Figure 5, the reconstruction phase extracts only the central region of each patch (25% of the patch) that has enough contextual information for effective feature extraction.

We also prefer our model to offer increased attention at the patch’s central region for superior results. Therefore, we introduce mask-PSNR, a novel loss function that utilizes a mask \((M)\) to let the network provide additional attention to the patch’s central region. As shown in the Figure 5, \(k\) depicts the size of the region that the model emphasizes the most. Mathematically, our loss function can be expressed as

\[ \text{Mask-PSNR} = 10 \log_{10} \frac{\sum_{u,v} M(u,v) [I_{\text{HR}}(u,v) - I_{\text{SR}}(u,v)]^2}{N^2 I_{\text{max}}^2}, \]  

(10)

where \(I_{\text{max}}\), \(N^2\) represent the maximum pixel intensity and size of the patch. \(M(u,v)\), \(I_{\text{HR}}(u,v)\), and \(I_{\text{SR}}(u,v)\) represent the value of mask, HR patch, and SR patch, at location \((u,v)\). We term the type of mask used in our work as “2-D box linear decay”.

D. Implementation Details

In our proposed algorithm, we utilize zero padding to maintain the output size of every convolution layer constant. Every Conv operation in the RDNLBs has \(G\) filters; however, shallow feature extraction, local, and global feature blending Conv layers have \(G_b = 64\) filters. For up-scaling (US), we utilize the
IV. SETTINGS

Dataset, Degradation Model, and Metrics: The digital ortho maps (DOMs) utilized in this work are captured from Kaguya terrain camera [18]. They span a range of \([-63^\circ, 63^\circ]\) in latitude and \([30^\circ, 33^\circ]\) in longitude. We generate both the input LR (100 m/pixel) and ground truth HR images (50 m/pixel) by applying the bicubic interpolation on the original DOMs (\(\sim 7.4\) m/pixel). The total images considered for training and testing are 120 and 21, respectively. The HR and LR images have a size of \(1820 \times 1820\) and \(910 \times 910\). The SR results are evaluated using PSNR and PSNRB [20].

Training Settings: We incorporate the following settings for lunar DOMs. We train our model for 75 epochs with a learning rate of \(10^{-4}\) and utilized Adam optimizer [21] to update the weights. Each LR and HR image is divided into overlapping patches of size \(48 \times 48\) and \(96 \times 96\), respectively, with an overlap region of 50%.

V. RESULTS

We compare our network with one of the state-of-the-art natural image SR model, RDN [11], trained on Kaguya DOMs. Table I illustrates the quantitative comparisons for various settings between RDN and our model (RDNLA) for \(\times 2\) SR. All the experiments in Table I follow the patch-based prediction approach. However, experiments that do not follow overlap reconstruction utilize the non-overlapping stacking approach to reconstruct the SR output.

The networks utilizing the L1 loss function depict the overlap reconstruction’s significance. Quantitatively, it demonstrates that the overlap reconstruction substantially alleviates the blocking artifacts in the models’ SR output, which can also be seen in Figure 4. As shown in Table I, PSNRB increases significantly by an average of \(\sim 0.14\) dB. Our loss function improves the PSNR and PSNRB by \(0.054\) dB and \(0.039\) dB (compare row 4,6).

Table II shows a comparison of RDN-I, RDN-P and RDNLA based on several parameters. All the models are evaluated on RTX-5000 NVIDIA GPU. The effect of the input size can be understood by comparing “GPU memory per prediction” and “peak instantaneous power” of RDN-
I and RDN-P. The memory occupied and power consumed decrease by 49% and 70%, respectively. However, the model’s SR performance (PSNR, PSNR-B) decreases, which is not preferred. Our proposed SR algorithm (RDNLA) provides better SR performance than the other models, consumes 48% lesser memory and 67% lesser power than the baseline, RDN-I.

The power available for satellite subsystems depends on various factors such as the type of satellite mission, number of subsystems and type of solar panels. For example, a solar panel mounted on Cubesat [22] has 7 W rating, commercial LEO satellites [23] consume 40 W, whereas LRO spacecraft power is 685 W and Selene’s main orbiter’s power is 3.5 kW [18].

Our proposed SR algorithm’s instantaneous power consumption can be increased/decreased by increasing/decreasing the patch size, making it executable even on satellites with lower power availability. Though our algorithm provides enhanced results while consuming less power and memory, the “SR image reconstruction time” is significantly increased due to an increase in patches, as shown in Table II. Future research could address this problem by proposing real-time SR algorithms that provide enhanced SR results and are executable on low-powered devices.

VI. ABLATION INVESTIGATION

Performance on natural images. We also train our network on the DIV2K dataset [24] to observe our algorithm’s performance on natural images. We follow the exact training settings as mentioned by [11]. We trained our model for a total of 800 epochs and utilize P = 16, D = 8, and G = 64 to report our results. We also employ a self-ensemble strategy [11] to boost the performance of RDNLA (RDNLA+). Table III shows the quantitative comparisons for x2, x3, and x4 SR on the Set5 [25], Set14 [26], and B100 [27] datasets. It can be observed that our proposed algo. outperforms all low-powered device SR models significantly but provides slightly lower performance than standard SR models. This may be because the proposed loss function lays little emphasis on the image’s boundary regions.

In Figures 6, 7, and 8 we show the visual comparisons on scale ×4. For images “119082” (Figure 6) and “302008” (Figure 8) we can observe that most of the compared methods generate hallucinated edges (high-frequency information) which is quite common in SISR due to limited information provision and training bias [11]. In contrast to the previous methods, RDNLA and RDNLA+ do not generate hallucinated information and thus, provides more reliable results proving effective for super-resolution of remote-sensing data. For image “005” (Figure 7) it can be clearly stated that the proposed algorithm retains and generates outputs possessing significantly more high-frequency information compared to the previous methods. The higher-frequency retention is possible due to mask PSNR, the non-local attention mechanism which helps increase the receptive field, and utilization of dense hierarchical features.

Study of CM, LRA, and GFB. The ablation study of the effects of coupled memory (CM), local residual attention (LRA), and global feature blending (GFB) has been depicted in Table IV. We utilize the DIV2K dataset to observe the effects significantly. In all the eight scenarios, there are a total of P = 16 RDNLBs and D = 6 Conv layers, with a growth rate of G = 32. The baseline network obtained without CM, LRA, and GFB (denoted as CM0LRA0GFB0) performs poorly with a PSNR of 36.39 dB. It indicates that non-overlapping stacking of convolution blocks does not result in a satisfactory output in deep networks. We then examined how each component contributes towards enhancing the SR output, which resulted in networks CM1LRA0GFB0, CM0LRA1GFB0, and CM0LRA0GFB1 (Col 2 – 4 in Table IV). Every component significantly boosts the model’s performance since it contributes to efficient information flow and hierarchical feature extraction. Similar effects can be observed when two and all
TABLE III
QUANTITATIVE RESULTS (PSNR/SSIM) ON NATURAL IMAGE SR BENCHMARK DATASETS. BEST AND SECOND BEST RESULTS ARE HIGHLIGHTED AND UNDERLINED.

| Method       | Set5      |              |              | Set14      |              |              | B100       |              |              |
|--------------|-----------|--------------|--------------|------------|--------------|--------------|------------|--------------|--------------|
|              | x2        | x3          | x4          | x2         | x3          | x4          | B100       | x2           | x3           |
| RDN [11]     | 38.24/0.9014 | 34.74/0.9298 | 32.47/0.8990 | 34.01/0.9212 | 30.65/0.8486 | 28.81/0.7671 | 32.34/0.9017 | 29.26/0.8093 | 27.72/0.7419 |
| RCAN [28]    | 38.27/0.9014 | 34.74/0.9299 | 32.63/0.9002 | 34.12/0.9216 | 30.65/0.8482 | 28.87/0.7889 | 32.41/0.9027 | 29.32/0.8111 | 27.77/0.7436 |
| SRFRN [20]   | 38.31/0.9009 | 34.76/0.9292 | 32.47/0.8983 | 33.82/0.9106 | 30.51/0.8461 | 28.81/0.7868 | 32.29/0.9010 | 29.24/0.8084 | 27.72/0.7409 |
| RANAN [10]   | 38.17/0.9011 | -            | 32.49/0.8982 | 33.87/0.9207 | -            | 28.81/0.7878 | 32.32/0.9014 | -            | 27.72/0.7421 |
| OSIR [31]    | 38.21/0.9012 | 34.72/0.9297 | 32.53/0.8992 | 33.94/0.9206 | 30.57/0.8470 | 28.86/0.7878 | 32.36/0.9019 | 29.29/0.8103 | 27.75/0.7428 |
| SAN [32]     | 38.31/0.9020 | 34.75/0.9300 | 32.64/0.9003 | 34.07/0.9213 | 30.59/0.8476 | 28.92/0.7888 | 32.42/0.9028 | 29.33/0.8112 | 27.78/0.7436 |

Low Powered Device SR

| Method       | Set5      |              |              | Set14      |              |              | B100       |              |              |
|--------------|-----------|--------------|--------------|------------|--------------|--------------|------------|--------------|--------------|
|              | x2        | x3          | x4          | x2         | x3          | x4          | B100       | x2           | x3           |
| MobileSR [18]| -         | -            | -            | 31.74/0.8843 | -            | -            | -          | -            | -            |
| Kim et al. [14]| 36.66/0.9548 | -          | -            | 32.52/0.9073 | -            | -            | -          | -            | -            |
| SRFRN [21]   | 37.06/0.9565 | 32.62/0.9099 | 31.47/0.8893 | 33.59/0.9258 | 29.74/0.8399 | 28.92/0.8013 | 30.41/0.8578 | 27.37/0.7534 | 26.86/0.7079 |
| SpleSR [10]  | -         | -            | 31.76/0.8902 | -            | -            | -            | -          | -            | -            |
| MFAGAN [10]  | -         | -            | 30.16/−      | -            | -            | -            | -          | -            | -            |
| RDNLA        | 38.12/0.9011 | 34.54/0.9286 | 32.25/0.8963 | 33.76/0.9196 | 30.42/0.8454 | 28.68/0.7840 | 32.27/0.9010 | 29.17/0.8080 | 27.63/0.7384 |
| RDNLA+       | 38.22/0.9015 | 34.68/0.9298 | 32.46/0.8987 | 33.86/0.9205 | 30.56/0.8469 | 28.82/0.7868 | 32.33/0.9017 | 29.26/0.8095 | 27.72/0.7404 |

Fig. 6. Visual results (4×) with model trained on DIV2K [24]. The results are presented on image “119082” from B100 [27] dataset.

Fig. 7. Visual results (4×) with model trained on DIV2K [24]. The results are presented on image “005” from Set14 [26] dataset.

TABLE IV
ABLATION INVESTIGATION ON THE EFFECTS OF CM, LRA, AND GFB. THE RESULTS ARE BASED ON SET5 [25] DATASET WITH A SCALING FACTOR X2 IN 75 EPOCHS.

|                | CM      | LRA     | GFB     | PSNR    |
|----------------|---------|---------|---------|---------|
| Different     |         |         |         |         |
| combinations  |         |         |         |         |
| of CM, LRA,  |         |         |         |         |
| GFB           |         |         |         |         |
| PSNR          | 36.39   | 36.95   | 36.45   | 36.81   |
|               | 37.07   | 37.35   | 37.06   | 37.34   |

the three components are utilized in the network (Col 5 – 8 in Table IV), providing further enhanced results with the best PSNR = 37.34 dB (CM1LRA1GFB1). It is observed that CM plays a critical role among the three, followed by GFB in image SR.

Study of D and P. We investigate the basic network parameters: the number of RDNLBs (P) and the number of Conv layers per RDNLB (D). As shown in Figure 9, larger P and D would lead to higher performance. This is mainly because a larger P and D increase the depth of the network which allows it to capture long-range dependent hierarchical features for higher performance. With the help of this study, we chose P = 16 and D = 6 for all lunar SR experiments.

Effect of Mask Size: Table V shows the results of the effect
Fig. 8. Visual results (4×) with model trained on DIV2K [24]. The results are presented on image “302008” from B100 [27] dataset.

Fig. 9. Convergence Analysis of RDNLA with different values of D and P of mask size, k, on the SR outputs. SR outputs of all the networks are reconstructed using an overlap reconstruction mechanism. The effective patch we extract during the overlap reconstruction is of dimension 48 × 48. However, as observed from Table V the best result is obtained at k = 54, with the reason being that at k = 48, the mask utilized does not provide enough attention to the contextual information at the boundary of the effective patch. Therefore, the model does not fully extract information at the boundaries and does not give the highest quality output.

**Table V**

| Loss     | k  | RDN      | Proposed |
|----------|----|----------|----------|
| L1       |    | 69.900/69.824 | 69.917/69.840 |
| Mask-PSNR| 48 | 69.916/69.831 | 69.951/69.897 |
|          | 54 | 69.934/69.853 | 69.971/69.879 |
|          | 58 | 69.921/69.843 | 69.965/69.889 |

The results demonstrate that irrespective of the mask size and architecture, networks trained with mask-PSNR give better SR results than the networks trained with L1 loss. Extensive analysis of the test results depicts that not only in terms of average values, but also for each and every reconstructed image, models trained with mask PSNR give at least similar or better outputs (higher PSNR and PSNRB) than models trained with L1 loss.

**Effect of Location:** We then plot the Center Latitude (vs) PSNR to evaluate the effect of DOMs’ location on the SR output. As shown in Figure 10, it is observed that the images at the equatorial region (around 0°) give significantly enhanced SR outputs compared to near-polar regions. It may be caused
by the simple cylindrical projection of the moon’s surface. Distortion of shape and scale in a cylindrical projection is minimal in equatorial regions and maximal at the polar regions.

VII. CONCLUSION

In this paper, we identify an unresolved issue of data utilization efficiency in satellite missions. We solve it by presenting a simple yet straightforward system design for selective data transmission based on in-orbit inferences. We include an in-orbit SR block in the system design to make better in-orbit inferences. We introduce a novel SR evaluation procedure, an overlap reconstruction mechanism, a new loss function (mask PSNR), and a residual dense non-local attention network for the SR algorithm executability on low-powered device platforms and enhance SR results. Experiments on lunar data showed that the proposed algorithm achieves better PSNR and PSNRB over the baseline residual dense network (RDN). Moreover, our proposed SR algorithm consumes 48% lesser memory and 67% lesser power than the RDN, making it more suitable for executing on low-powered devices.

REFERENCES

[1] M. R. Arefin, V. Michalski, P.-L. St-Charles, A. Kalaitzis, S. Kim, S. E. Kahou, and Y. Bengio, “Multi-image super-resolution for remote sensing using deep recurrent networks,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, June 2020.

[2] Z. Bing, “Intelligent remote sensing satellite system,” Journal of Remote Sensing, vol. 15, no. 3, pp. 415–431, 2011.

[3] L. Liebel and M. Körner, “Single-image super resolution for multispectral remote sensing data using convolutional neural networks,” ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 41, pp. 883–890, 2016.

[4] S. Lei, Z. Shi, and Z. Zou, “Super-resolution for remote sensing images via local–global combined network,” IEEE Geoscience and Remote Sensing Letters, vol. 14, no. 8, pp. 1243–1247, 2017.

[5] D. Zhang, J. Shao, X. Li, and H. T. Shen, “Remote sensing image super-resolution via mixed high-order attention network,” IEEE Transactions on Geoscience and Remote Sensing, pp. 1–14, 2020.

[6] X. Dong, L. Wang, X. Sun, X. Jia, L. Gao, and B. Zhang, “Remote sensing image super-resolution using second-order multi-scale networks,” IEEE Transactions on Geoscience and Remote Sensing, 2020.

[7] Y. Kim, J.-S. Choi, and M. Kim, “A real-time convolutional neural network for super-resolution on fpga with applications to 4k uhd 60 fps video services,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 8, pp. 1925–1934, 2018.

[8] R. Lee, S. I. Venieris, L. Dudiżak, S. Bhattacharya, and N. D. Lane, “Mobisr: Efficient on-device super-resolution through heterogeneous mobile processors,” in The 25th Annual International Conference on Mobile Computing and Networking, 2019, pp. 1–16.

[9] X. Liu, Y. Li, J. Fromm, Y. Wang, Z. Jiang, A. Mariakakis, and S. Patel, “SplitSR: An end-to-end approach to super-resolution on mobile devices,” Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, vol. 5, no. 1, pp. 1–20, 2021.

[10] W. Cheng, M. Zhao, Z. Ye, and S. Gu, “MFAGAN: A compression framework for memory-efficient on-device super-resolution gan,” arXiv preprint arXiv:2107.12674, 2021.

[11] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, “Residual dense network for image restoration,” IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1–1, 2020.

[12] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 2017.

[13] X. Dong, Z. Xi, X. Sun, and L. Gao, “Transferred multi-perception attention networks for remote sensing image super-resolution,” Remote Sensing, vol. 11, no. 23, 2019.

[14] Z. Pan, W. Ma, J. Guo, and B. Lei, “Super-resolution of single remote sensing image based on residual dense backprojection networks,” IEEE Transactions on Geoscience and Remote Sensing, vol. 57, no. 10, pp. 7918–7933, 2019.

[15] X. Glorot, A. Bordes, and Y. Bengio, “Deep sparse rectifier neural networks,” in Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, ser. Proceedings of Machine Learning Research, G. Gordon, D. Dunson, and M. Dudík, Eds., vol. 15. Fort Lauderdale, FL, USA: JMLR Workshop and Conference Proceedings, 11–13 Apr 2011, pp. 315–323.

[16] Y. Tai, J. Yang, X. Liu, and C. Xu, “Mmnet: A persistent memory network for image restoraton,” in Proceedings of the IEEE International Conference on Computer Vision (ICCV), Oct 2017.

[17] X. Wang, R. Girshick, A. Gupta, and K. He, “Non-local neural networks,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 7794–7803.

[18] J. Haruyama, T. Matsunaga, M. Ohtake, T. Morota, C. Honda, Y. Yokota, M. Torii, and Y. Ogawa, “Global lunar-surface mapping experiment using the lunar imager/spectrometer on selene,” Earth, planets and space, vol. 60, no. 4, pp. 243–255, 2008.

[19] C. Chen, P. Zhang, H. Zhang, J. Dai, Y. Yi, H. Zhang, and Y. Zhang, “Deep learning on computational-resource-limited platforms: A survey,” Mobile Information Systems, 2020.

[20] C. Yin and A. C. Bovik, “Quality assessment of deblocked images,” IEEE Transactions on Image Processing, vol. 20, no. 1, pp. 88–98, 2011.

[21] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” ICLR, 2014.

[22] J. Claricots and S. M. Dakka, “Design of power, propulsion, and thermal sub-systems for a 3u casbus measuring earth’s radiation imbalance,” Aerospace, vol. 5, no. 2, p. 63, 2018.

[23] A. Kharsany, “Power modeling and budgeting design and validation with in-orbit data of two commercial leo satellites,” 2017.

[24] R. Timofte, E. Agustsson, L. Van Gool, M.-H. Yang, and L. Zhang, “Ntire 2017 challenge on single image super-resolution: Methods and results,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, July 2017.

[25] A. Bevilacqua, A. Roumy, C. Guillenmot, and M. L. Alberi-Morel, “Low-complexity single-image super-resolution based on nonnegative neighbor embedding,” 2012.

[26] R. Zeyde, M. Elad, and M. Protter, “On single image scale-up using sparse-representations,” in Curves and Surfaces, J.-D. Boissonnat, P. Chenin, A. Cohen, C. Gout, T. Lyche, M.-L. Mazure, and L. Schumaker, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 711–730.

[27] D. Martin, C. Fowlkes, D. Tal, and J. Malik, “A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics,” in Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001, vol. 2, 2001, pp. 416–423 vol.2.

[28] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, “Image super-resolution using very deep residual channel attention networks,” in Proceedings of the European Conference on Computer Vision (ECCV), September 2018.

[29] Z. Li, J. Yang, Z. Liu, X. Yang, G. Jeon, and W. Wu, “Feedback network for image super-resolution,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2019.

[30] Y. Zhang, K. Li, K. Li, B. Zhong, and Y. Fu, “Residual non-local attention networks for image restoration,” in International Conference on Learning Representations, 2019.

[31] X. He, Z. Mo, P. Wang, Y. Liu, M. Yang, and J. Cheng, “Ode-inspired network design for single image super-resolution,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2019.

[32] T. Dai, J. Cai, Y. Zhang, S.-T. Xia, and L. Zhang, “Second-order attention network for single image super-resolution,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2019.

[33] J. Lee, J. Lee, and H.-J. Yoo, “SRNPU: An energy-efficient cnn-based super-resolution processor with tile-based selective super-resolution in mobile devices,” IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol. 10, no. 3, pp. 320–334, 2020.

[34] C. Dong, C. C. Loy, K. He, and X. Tang, “Image super-resolution using deep convolutional networks,” IEEE transactions on pattern analysis and machine intelligence, vol. 38, no. 2, pp. 295–307, 2015.

[35] J. Kim, J. K. Lee, and K. M. Lee, “Deeply-recursive convolutional network for image super-resolution,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2016.
[36] J. Kim, J. Kwon Lee, and K. Mu Lee, “Accurate image super-resolution using very deep convolutional networks,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2016.

[37] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang, “Deep laplacian pyramid networks for fast and accurate super-resolution,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 624–632.

[38] Y. Tai, J. Yang, and X. Liu, “Image super-resolution via deep recursive residual network,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 2017.

[39] M. Haris, G. Shakhnarovich, and N. Ukita, “Deep back-projection networks for super-resolution,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2018.