



Super-resolution reconstruction of UAV-borne gamma-ray spectrum images based on Real-ESRGAN algorithm

Xin Wang¹ · Yuan Yuan¹ · Xuan Zhao¹ · Guang-Hao Luo¹ · Qi-Qiao Wei¹ · He-Xi Wu¹ · Chao Xiong¹

Received: 9 July 2024 / Revised: 14 September 2024 / Accepted: 9 October 2024 / Published online: 4 January 2026

© The Author(s), under exclusive licence to China Science Publishing & Media Ltd. (Science Press), Shanghai Institute of Applied Physics, the Chinese Academy of Sciences, Chinese Nuclear Society 2025

Abstract

Unmanned aerial vehicle (UAV)-borne gamma-ray spectrum survey plays a crucial role in geological mapping, radioactive mineral exploration, and environmental monitoring. However, raw data are often compromised by flight and instrument background noise, as well as detector resolution limitations, which affect the accuracy of geological interpretations. This study aims to explore the application of the Real-ESRGAN algorithm in the super-resolution reconstruction of UAV-borne gamma-ray spectrum images to enhance spatial resolution and the quality of geological feature visualization. We conducted super-resolution reconstruction experiments with 2×, 4× and 6× magnification using the Real-ESRGAN algorithm, comparing the results with three other mainstream algorithms (SRCNN, SRGAN, FSRCNN) to verify the superiority in image quality. The experimental results indicate that Real-ESRGAN achieved a structural similarity index (SSIM) value of 0.950 at 2× magnification, significantly higher than the other algorithms, demonstrating its advantage in detail preservation. Furthermore, Real-ESRGAN effectively reduced ringing and overshoot artifacts, enhancing the clarity of geological structures and mineral deposit sites, thus providing high-quality visual information for geological exploration.

Keywords UAV-borne gamma-ray spectrum · Super-resolution reconstruction · Real-ESRGAN · Image processing

1 Introduction

UAV-borne gamma-ray spectrum survey is to measure and record the type and intensity of gamma-ray emitted by natural radionuclides (e.g., uranium, thorium, potassium) on the ground and shallow surfaces. These data, combined with the coordinate information, are processed, screened and corrected to produce maps of surface radionuclide distribution, essential for geological exploration, radioactive mineral prospecting, and environmental monitoring. Compared to fixed-wing aircraft, UAV-borne gamma-ray spectrum surveys offer low flight costs, high measurement efficiency, safe and flexible flight processes, and prolonged hovering

capabilities. Consequently, they have been widely adopted in radioactive mineral exploration, radiation environment monitoring, nuclear emergency response, and related fields [1–3].

Zhang [4] skillfully used tension spline method and Munsell transform technology to carry out fine rasterization and anomaly correction of aerial gamma spectrum data, successfully extracted key information, clearly displayed the energy spectrum characteristics of Xiangshan volcanic basin, and analyzed the distribution of uranium, thorium and potassium, providing new clues for uranium exploration and promoting the development of geological exploration technology. Šálek [5] tested the performance of a new type of small airborne gamma-ray spectrum measurement equipment based on UAV and verified its ability to detect uranium anomalies at different altitudes and can accurately capture small changes in gamma-ray intensity even at higher altitude. Li [6] developed a curviline-based noise reduction technique for airborne gamma data, which can reconstruct multi-scale data for different analysis purposes, effectively remove noise from the data, while retaining local anomaly information to avoid

This work was supported by the National Natural Science Foundation of China (Nos. 12205044 and 12265003) and 2024 Jiangxi Province Civil-Military Integration Research Institute ‘BeiDou+’ Project Subtopic (No. 2024JXRH0Y06).

✉ Chao Xiong
xiongchaollg@ecut.edu.cn

¹ School of Nuclear Science and Engineering, East China University of Technology, Nanchang 330013, China

resolution loss and boundary effects. Lyu [7] proposed a K-factor prediction model considering flight altitude and direct distance. The model pays special attention to the key factors such as path loss, shadow fading and small-scale fading at different flight altitudes, which brings potential application value to the field of aerial gamma-ray spectral image processing. Wang [8] proposes a layered approach specifically for REM recovery of limited sampled data in unknown environments, demonstrating the possibility of achieving high-precision REM construction under low sampling rates. The data recovery algorithm and sampling optimization strategy can be used for reference in the research of aerial gamma-ray spectral image processing to improve the processing accuracy and efficiency.

The demand for high-resolution UAV-borne gamma-ray spectrum images has increased to improve the accuracy of research. For instance, in sandstone-type uranium exploration and radiation environment monitoring, high-resolution images allow researchers to identify weak anomalies and locate radionuclide distributions more accurately. However, the direct acquisition of ultra-high-resolution gamma-ray spectrum images is challenging due to UAV load and flight altitude limitations.

Image super-resolution (SR) technology, which converts low-resolution (LR) images into high-resolution (HR) images, can significantly enhance the details and data information of existing images. This technology plays an important role in many fields such as satellite imagery [9], face recognition [10], and medical imaging [11]. In recent years, with the development of deep learning, people begin to use deep learning technology to improve image super-resolution. The SR method based on convolutional neural network, residual network (ResNet) and generative adversarial networks (GAN) are proposed. SRCNN [12], as the first SR deep learning network, achieves fast online application with its lightweight structure and excellent recovery quality, despite its large computational load. ESPCN [13] introduces an innovative subpixel convolution layer to achieve LR to HR super-resolution at a minimal computational cost. VDSR [14] uses a very deep convolutional network, combined with residual learning, to significantly improve the convergence speed during training. These methods improve the accuracy and speed of image super-resolution by using faster and deeper convolutional neural networks. However, when enlarging the LR image to the HR image, the reconstructed SR image often lacks texture details due to the large-scale factor, resulting in unsatisfactory reconstruction effect. SRGAN [15–17] enhanced the content loss function with counter-loss by training GAN, and replaced the content loss based on mean squared error (MSE) with the loss based on VGG network feature map, effectively overcoming the problem of low perception quality of reconstructed images and making the generated images more realistic. However,

the downside of this method lies in its complex network structure and lengthy training process.

However, degradation of image quality is a complex phenomenon, usually due to the limitations of the imaging system and environmental disturbances. Blind super-resolution (blind SR) technology arises at the historic moment. It can effectively restore LR images to HR images when the degradation process is not completely clear. This technology can be divided into two categories: display modeling and implicit modeling. Display modeling by parameterizing fuzzy kernel and noise information. SRMD [18–20], as the first deep learning-driven blind SR method, introduces a dimension expansion strategy, which enables the convolution network to handle fuzzy kernel and noise level inputs, and solves the dimension mismatch problem, although it may perform poorly for uncovered degradation types. The BSRGAN model [21] proposed by Zhang et al. in 2021 enhances the adaptability of models to real-world degradation by introducing complex degradation factors and random shuffling strategies. Based on the kernel stretching strategy of Luo [22] proposed a new practical degradation model, which uses dynamic depth linear filter and constrained least square deconvolution algorithm based on neural network to improve the restoration quality of blurred images. Implicit modeling abandons explicit parameters and uses deep learning techniques, especially GAN, to restore directly from LR images to HR images. The CinCGANcite [23, 24] model adopts double cyclic GAN structure, which effectively solves the problem of complex interference in low-resolution input. The research classification of Liu [25] points out that some methods train SR models by learning the degradation process from HR to LR and using the generated LR samples, such as Degradation GAN [26], FSSR [27] and FSSRGAN [28], but these methods may have domain gap problems. DASR improves the performance of SR training through domain gap awareness training and domain distance weighted monitoring strategy. Real-ESRGAN [29] improved on the basis of ESRGAN [30], using discriminator designed by U-Net and high-order degradation process, and introducing sinc filter to reduce ringing and overshoot artifacts, providing a more accurate and stable super-resolution solution for real-world images.

In this paper, a new scheme combining UAV-borne gamma-ray spectrum survey and image super-resolution technology is proposed to overcome the limitations of existing technology in obtaining high-resolution images. Real-ESRGAN technology can improve the clarity and feature enhancement of UAV-borne gamma-ray spectrum images and then help to interpret geological data more accurately. For example, it can enhance areas of the image that are blurred due to topographic effects, making geological features more clearly visible. In addition, Real-ESRGAN's de-noising and detail enhancement capabilities help to

remove noise generated during flight measurements, thereby improving data quality. Image super-resolution technology can improve LR image to HR through algorithm processing, and compensate for the shortage of hardware equipment. This technology, combined with the UAV-borne gamma-ray spectrum survey, can not only improve the clarity and detail expression of the image, but also effectively improve the accuracy and reliability of data analysis. Therefore, the combination of UAV-borne gamma-ray spectrum survey technology and image super-resolution technology can more effectively monitor environmental radiation levels, but also provide more accurate data support for the efficient exploration of radioactive minerals, and promote the development of radioactive geophysical exploration technology.

2 Technical principles and methods

2.1 Technical principles and methods

UAV-borne gamma-ray spectrum technology captures gamma rays of different energy levels from the ground and shallow surfaces using a gamma detector mounted on the UAV. This technology is promising for geological exploration and radiation environment monitoring. However, the images generated often suffer from noise and disturbances, leading to potential misinterpretations of geological information.

In the field of image processing, a series of noise filtering and image enhancement techniques are usually needed to improve the quality of UAV-borne gamma-ray spectrum images. The Gaussian filter is widely used, which can reduce the interference of noise to the image quality by smoothing the image. The formula of Gaussian filtering [31] can be expressed as:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \quad (1)$$

where $G(x, y)$ is a Gaussian kernel function, and in practical applications, this kernel function is usually discretized and applied to an image. x and y are the positions of pixels in the image. σ is the standard deviation of the Gaussian distribution. However, when the Gaussian filter is used for global smoothing of the whole image, it lacks the ability of self-adaptation and cannot smooth in different degrees according to the local features of the image.

Median filtering [32] is also a common nonlinear filtering technique, which reduces noise by replacing the value of a pixel with the median value in its neighborhood. For a given image I , the formula for calculating the pixel value of the image I' after median filtering at (x, y) is:

$$I'(x, y) = \text{median}\{I(x+i, y+j) | (i, j) \in W\}, \quad (2)$$

where $I(x, y)$ is the pixel value of the original image at the position (x, y) . $I'(x, y)$ is filtered image at position (x, y) of the pixel values. W is a neighborhood window, usually a $(2k+1) \times (2k+1)$ rectangular window with center (x, y) . median represents the median value of the pixel value in the neighborhood window.

Although this method surpasses Gaussian filtering in preserving edge information, it may cause boundary blurring when processing images with high contrast edges. To further emphasize geological features in the AGS images, image enhancement techniques such as contrast enhancement are used to improve the dynamic range of the images, making subtle details more visible. Edge detection technology can identify boundary lines in the image, which is particularly critical for determining the precise location of geological structures. However, these techniques may increase the artificial distortion of images, leading to over-emphasis or misidentification of geological features.

2.2 Image super-resolution reconstruction

Image super-resolution aims to convert LR images into HR images by modeling the degradation process, which includes blurring, downsampling, and noise. The goal is to find an operator that makes the reconstructed image as close as possible to the original HR image. Under the Bayesian framework, this process is expressed as a maximum a posteriori (MAP) problem [33], which involves a least squares problem with a regularization term to incorporate prior knowledge. The core goal of image super-resolution reconstruction technology is to convert LR images into HR images. Mathematically, this process is often modeled as a degradation process where the LR image is regarded as the result of the HR image being blurred, downsampled, and disturbed by noise.

$$I_{LR} = \mathbf{H} \cdot I_{HR} + n, \quad (3)$$

where \mathbf{H} is a degenerate operator, which includes processes such as downsampling, blurring and noise, and can be expressed as a matrix or an integral operator. n stands for observation noise, which may include instrument noise, environmental noise, etc. The general production is assumed to be Gaussian distribution or Poisson distribution.

The goal of super-resolution reconstruction is to find an operator R , which makes $R(I_{LR})$ as close as possible to the original high-resolution image I_{HR} . The formula can be as follows:

$$I_{HR}^{\text{est}} = R(I_{LR}), \quad (4)$$

where I_{HR}^{est} is the estimated high-resolution image.

Under the Bayesian framework [34], super-resolution reconstruction can be expressed as a MAP problem.

$$R^* = \operatorname{argmax}_R p(I_{HR} | I_{LR}), \tag{5}$$

where R^* is the optimal reconstruction operator, $p(I_{HR} | I_{LR})$ is the posterior probability of I_{HR} given the observed I_{LR} .

In practice, since it is usually not feasible to compute $p(I_{HR} | I_{LR})$ directly, the super-resolution reconstruction algorithm usually uses a regularization term instead, which translates into a least squares problem:

$$R^* = \operatorname{argmin}_R \left\{ \|I_{LR} - \mathbf{H} \cdot R(I_{LR})\|_2^2 + \lambda \cdot \Omega(R(I_{HR})) \right\}, \tag{6}$$

where $\|\cdot\|_2$ represents the L2 norm and is used to measure the reconstruction error. $\Omega(\cdot)$ is a regularization term (such as gradient smoothing, sparse representation, etc.) used to introduce prior knowledge about I_{HR} . λ is a regularization parameter used to balance data fitting and the importance of regularization.

In the field of deep learning, Real-ESRGAN and other algorithms approximate the optimal reconstruction operator R^* by training a depth neural network. At the same time, the GAN is used to improve the visual quality and naturalness of the reconstructed image.

3 Super-resolution reconstruction of gamma-ray spectrum image of UAV

3.1 Real-ESRGAN principle

ESRGAN (enhanced super-resolution generative adversarial networks), based on GAN, generates high-quality super-resolution images through competition between generator and discriminator. Real-ESRGAN improves upon ESRGAN specifically for the super-resolution reconstruction of

real-world images. By introducing higher-order degradation models and Sinc filters, it can more accurately simulate the image degradation process in the real world, including blur, downsampling, noise and JPEG compression among others. Additionally, Real-ESRGAN employs a U-Net structure discriminator and spectral normalization technology, which not only enhances the discriminator’s ability to distinguish, but also improves the stability of the training process.

(1) GAN

GAN (generative adversarial networks) [35] consists of two main components: generator and discriminator as shown in Fig. 1. The generator is responsible for converting the low-resolution image into a high-resolution image, while the discriminator is responsible for distinguishing between the generated high-resolution image and the real high-resolution image. The two are trained through a process of adversarial interaction, where the generator continually refines its output to deceive the discriminator, while the discriminator concurrently enhances its capabilities to more effectively distinguish between generated and real images.

(2) Network architecture

Real-ESRGAN retains the residual-in-residual dense block (RRDB) [36] from ESRGAN as the core component of its generator. Through an innovative network structure design, the performance of image super-resolution reconstruction is significantly improved. The RRDB is composed of multiple residual blocks, each of which is further connected in a dense manner, allowing features from all previous layers to be directly connected to the current layer. RRDB enhances the learning capability of the network through multiple residual connections while simultaneously eschewing the use of batch normalization (BN) layers, which contributes to the improvement of the detail clarity in the generated images. In addition, the RRDB’s role in a generator is multifaceted. It not only acts as the backbone for feature extraction, but also works collaboratively with other parts of the network through residual connections, such as sampling

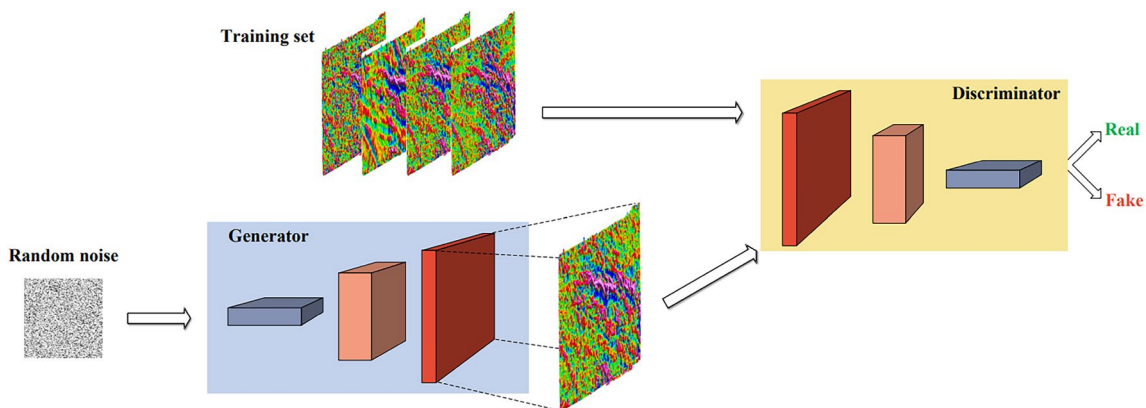


Fig. 1 (Color online) Basic architecture of GAN

layers and feature fusion modules, to generate high-quality high-resolution images. Figure 2 shows the structure of the RRDB. By this design, Real-ESRGAN is able to produce detailed and visually realistic super-resolution images to meet the needs of a variety of practical applications.

(3) Loss function

To achieve high-quality reconstruction, Real-ESRGAN employs a composite loss function, a design that ensures the accuracy of the generated high-resolution images at the pixel level as well as visual fidelity and richness of detail. The compound loss function consists of the following key components:

1) L1 loss (mean absolute error, MAE)

L1 loss measures the average of the absolute difference between the predicted value and the true value, and its mathematical expression is defined as:

$$L1 \text{ Loss} = \frac{1}{N} \sum_{i=1}^N |I_{HR}^{(i)} - \hat{I}_{HR}^{(i)}| \tag{7}$$

Including $I_{HR}^{(i)}$ is the i th a real high-resolution image pixel values, $\hat{I}_{HR}^{(i)}$ is corresponding to generate the image pixel value, and N is the total pixels.

2) Perceptual Loss

Perceptual loss is typically based on pre-trained convolutional neural networks (such as VGG networks) to extract features and compare the differences in these features between the generated and real images. A simplified perceived loss can be expressed as:

$$\text{Perceptual Loss} = \frac{1}{M} \sum_{j=1}^M \left\| F(I_{HR}^{(i)})_j - F(\hat{I}_{HR}^{(i)})_j \right\|_2, \tag{8}$$

where F represents the feature extraction network, M represents the total number of elements in the feature map, and $\| \cdot \|_2$ represents the Euclidean norm.

3) Adversarial Loss

By training a discriminator to distinguish between a real image and a generated image, the generator's goal is to maximize the probability that the discriminator will make a wrong judgment. A common form of countering loss is the use of *Wasserstein* distance, which is expressed as:

$$\text{Adversarial Loss} = - \mathbb{E}_{I_{HR} \sim P_{\text{data}}(I_{HR})} [D(I_{HR})] + \mathbb{E}_{\hat{I}_{HR} \sim P_h(\hat{I}_{HR})} [D(\hat{I}_{HR})], \tag{9}$$

where D is the discriminator and P_{data} and P_h are the distribution of the real image and the generated image, respectively.

The loss functions synergistically work together to ensure that the generated high-resolution images exhibit not only fidelity at the pixel level, but also visually realistic and detailed enough to meet the needs of a variety of real-world applications.

3.2 Data source

The data in this paper are from a experimental area in northern Gansu province, China, which is shown in Fig. 3. A total of 62889 data point from total count (Tc), uranium

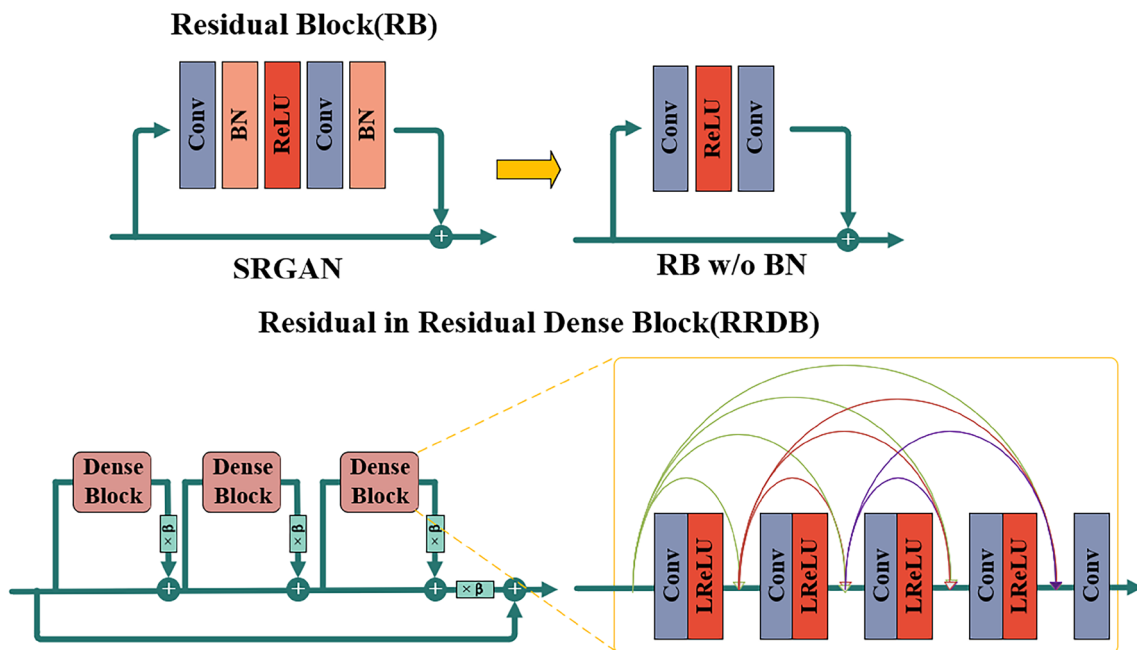


Fig. 2 (Color online) RRDB structure diagram

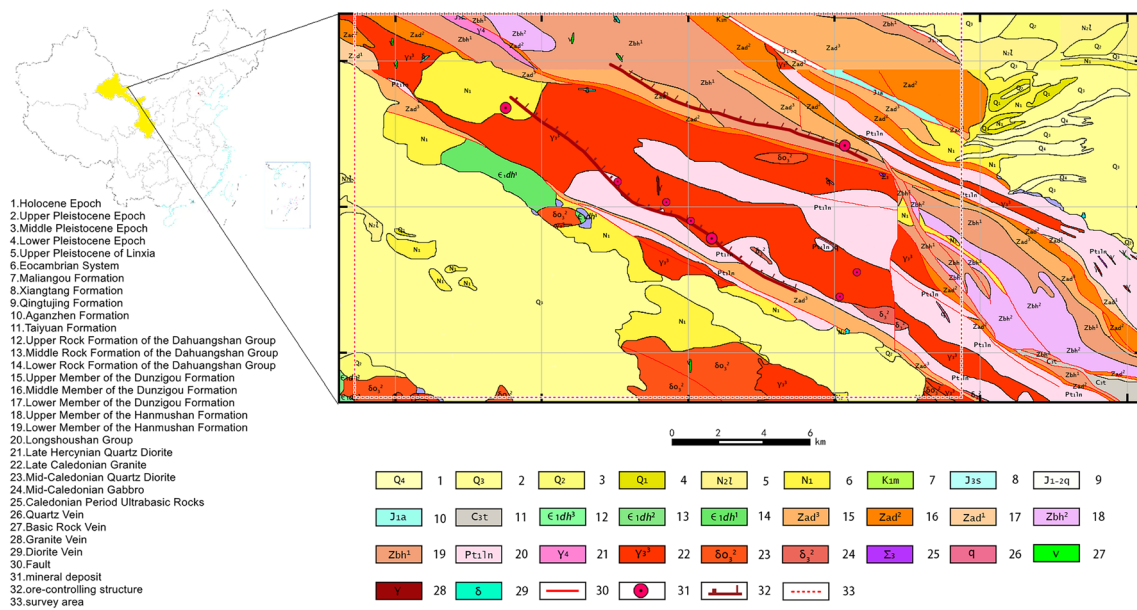


Fig. 3 (Color online) Geological map of the survey area

(U), thorium (Th) and potassium (K) channels were measured and processed to create equivalent maps of surface radionuclide distribution. To enhance the visual quality and detail resolution of UAV-borne gamma-ray spectrum images, image restoration techniques were employed on the raw data in this study. Consequently, the color scales in the UAV-borne gamma-ray spectrum images presented herein are dimensionless, serving solely to enhance contrast and visualization rather than for quantitative measurement of radioactive count rates. In our study, in order to adapt to the specific resolution and characteristics of the UAV-borne gamma-ray spectrum data, we performed segmentation and selection on different regions of the original UAV-borne gamma-ray spectrum images, directly skipping areas where the cropping region exceeds the image boundaries, ultimately forming 232 two-dimensional slices of 80×80 pixels each. The 80×80 pixel cutting size was chosen to take full advantage of the gamma-ray detector's spatial resolution at the actual flight altitude, while maintaining enough detail for effective geological feature analysis.

In order to obtain low-resolution UAV-borne gamma-ray spectrum images, the 'resize' function from the 'PIL Image' library was utilized to alter the image size by specifying new dimension parameters (width and height). The 'resample=Image.LANCZOS' parameter specifies the use of the 'Lanczos' resample algorithm, which is a high-quality resampling method suitable for image scaling. As shown in Fig. 4, with the increase in the UAV-borne gamma spectrum image to 6 times, some geological detail loss can be clearly observed.

3.3 Model training and analysis

The hardware used includes an INTEL I5-13600KF 14-Core processor with 3.5 GHz, 32 GB memory and an NVIDIA GeForce RTX 4070Ti graphics card. The RealESRGAN model was constructed using PyTorch framework, and the Adam optimizer was used to train the model at a learning rate of 1×10^{-4} , exponential motion average (EMA) was employed for more stable training. In addition, L1 loss, perceived loss and GAN loss were combined for training with weights of (1,1,0.1), respectively. L1 loss ensures the accuracy of the reconstructed image at the pixel level, perceptual loss is enhances the high-level visual quality of the image to align with human perception, and GAN loss increases the realism and naturalness of the image. The comprehensive use of different types of loss functions to improve the effect of UAV-borne gamma-ray spectrum super-resolution image reconstruction.

UAV-borne gamma-ray spectrum images are a special type of remote sensing image, providing important data for geological exploration, mineral resource development and environmental monitoring by detecting and recording the distribution of radioactive elements on the surface. These images need high spatial resolution to capture subtle geological features. The main advantage of super-resolution image is their ability to significantly enhance the spatial resolution of the image, making the visual effect of the image more clearer, detailed and more realistic. To evaluate the application of super-resolution reconstruction technology in UAV-borne gamma-ray spectrum images, this study selected the peak signal-to-noise ratio (PSNR) and

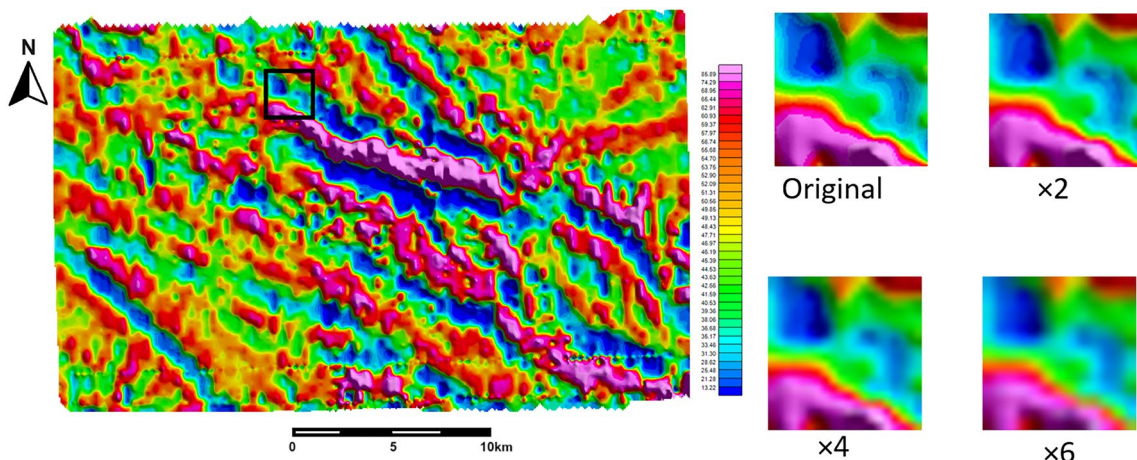


Fig. 4 (Color online) Comparison between image and original image at different multiples of resolution

structural similarity index (SSIM) as the main evaluation indicators.

PSNR is a crucial metric for measuring image quality, primarily used to evaluate the similarity between the reconstructed image and the original image, providing a quantitative analysis of pixel-level errors between images. Its unit is decibels (dB), a higher value, indicates that the reconstructed image is to the original image, thereby reflecting better image quality. The calculation of PSNR depends on MSE, the formula is:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (I(i, j) - K(i, j))^2, \tag{10}$$

where $I(i, j)$ is the pixel value of the original image, $K(i, j)$ is the pixel value of the reconstructed image, and m and n are the width and height of the image, respectively. The calculation formula of PSNR is:

$$PSNR = 10 \log_{10} \left(\frac{MAX^2}{MSE} \right), \tag{11}$$

where MAX represents the maximum possible value of an image pixel. If the pixel value is 8 bits (i.e., the range is 0–255), then MAX=255. The MAX corresponding to the image in this article is 255.

SSIM is an index used to measure the similarity between two images. It evaluates the image quality through brightness, contrast and structure to reflect the perceived quality of the image more comprehensively. Therefore, the calculation of SSIM is also based on three main components: brightness, contrast and structure. For the original image x and the reconstructed image y , the SSIM index is:

$$SSIM(x, y) = [l(x, y)]^\alpha [c(x, y)]^\beta [s(x, y)]^\gamma \tag{12}$$

where $l(x, y) = \frac{2\mu_x\mu_y+C_1}{\mu_x^2+\mu_y^2+C_1}$ is the brightness comparison, μ_x and μ_y are the average grays of the original image x and the reconstructed image y , C_1 is a constant that is not zero, $c(x, y) = \frac{2\sigma_x\sigma_y+C_2}{\sigma_x^2+\sigma_y^2+C_2}$ is contrast comparison, σ_x and σ_y are the standard deviations of the grayscale of the original image x and the reconstructed image y , C_2 is a constant that is not zero, $s(x, y) = \frac{\sigma_{xy}+C_3}{\sigma_x\sigma_y+C_3}$ is structural comparison, σ_{xy} is the covariance of the gray scale of the original image x and the reconstructed image y , C_3 is usually $C_2/2$. α , β and γ are weight parameters, usually taking the value 1. The final SSIM index formula is combined as follows:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}. \tag{13}$$

The value ranges from 0 to 1. The closer the SSIM value is to 1, the more similar the reconstructed image is to the original image and the higher the quality is. On the contrary, the lower the SSIM value, the greater the difference between the reconstructed image and the original image, and the worse the quality.

4 Results and discussion

In this paper, we reconstructed the UAV-borne gamma-ray spectrum images with 2×, 4× and 6× super-resolution. SRCNN algorithm, SRGAN algorithm, FSRCNN algorithm and Real-ESRGAN algorithm are compared to verify the effectiveness of Real-ESRGAN algorithm.

To validate the model performance, this study randomly selected 10 images from the image set of four elements: U, Th, K, and Tc, to constitute a test set. During the testing

process, Th and K were primarily chosen for analysis. By comparing and analyzing with the geological map of the survey area (as shown in Fig. 3), it was found that the distribution patterns of Th and K within the survey area exhibit a high degree of consistency. Notably, there is a fault within the survey area that extends from the northwest to the southeast, which aligns closely with the high-value bands of thorium and potassium. This phenomenon is attributed to the distribution of Hercynian late-stage granites, which have caused an anomaly enrichment of radioactive potassium and thorium in survey area. Typically, granites contain minerals such as potassium feldspar and mica, which are rich in potassium; concurrently, they also contain thorium-bearing minerals like monazite and xenotime. Consequently, in areas where granite is distributed, the content of thorium and potassium is generally high. Based on this, the test dataset employed in this paper can authentically reflect the distribution of natural radioactive nuclides within the survey area.

To visually demonstrate the superior effect of the Real-RESGAN algorithm on the super-resolution reconstruction of UAV-borne gamma-ray spectrum images, the three introduced super-resolution reconstruction algorithms were applied with 2×, 4× and 6× super-resolution on the test set. The results are shown in Figs. 5, 6, and 7.

The analyses delineated within Figs. 5, 6 and 7 expose discernible disparities in the efficacy of assorted algorithms engaged in the reconstruction of imagery. The super-resolution reconstructions yielded by the SRCNN algorithm exhibit a measure of obscurity and chromatic aberration, with a pronounced exacerbation in blurriness concomitant to escalated magnification. Particularly, at the 6× magnification exemplified in Fig. 7, the loss of detail information leads to the obscurity of geological body boundaries, rendering them difficult to discern. The FSRCNN algorithm, an advanced rendition of its progenitor, ameliorates the acuity of high-frequency detail capture through the refinement of convolutional strata and activation mechanisms, thereby bestowing a marked enhancement in the textural, marginal, and configurational fidelity of the reconstructed imagery relative to the SRCNN (as depicted in Fig. 5). Despite these advancements, discrepancies persist between the FSRCNN’s output and the authentic geological background and boundaries, with residual haziness observable (as illustrated in Fig. 7). The SRGAN algorithm surpasses its precursors in the holistic quality of image reconstruction, integrating the GAN architecture to facilitate a more nuanced rendition of intricate image details. Even with the magnification increased to 6×, the detailed information and boundaries of the geological bodies remain

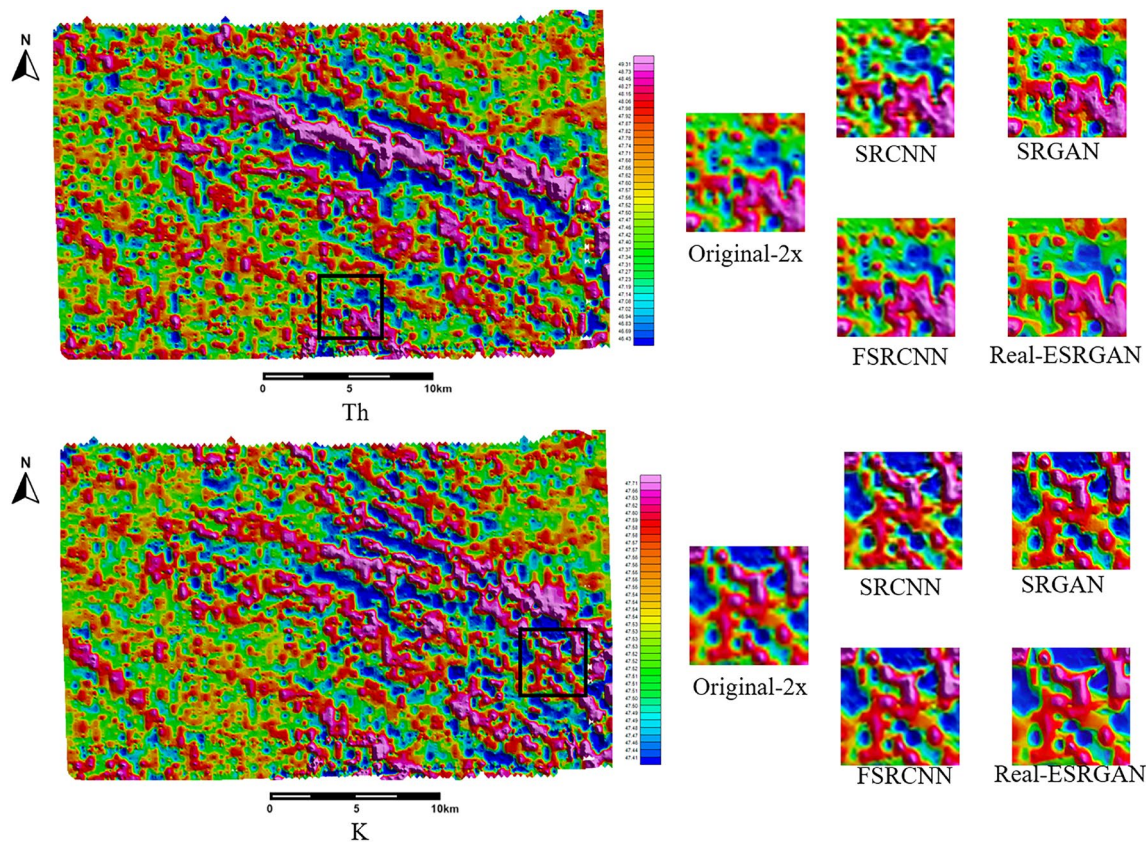


Fig. 5 (Color online) Comparison of results of 2× super-resolution reconstruction without algorithms

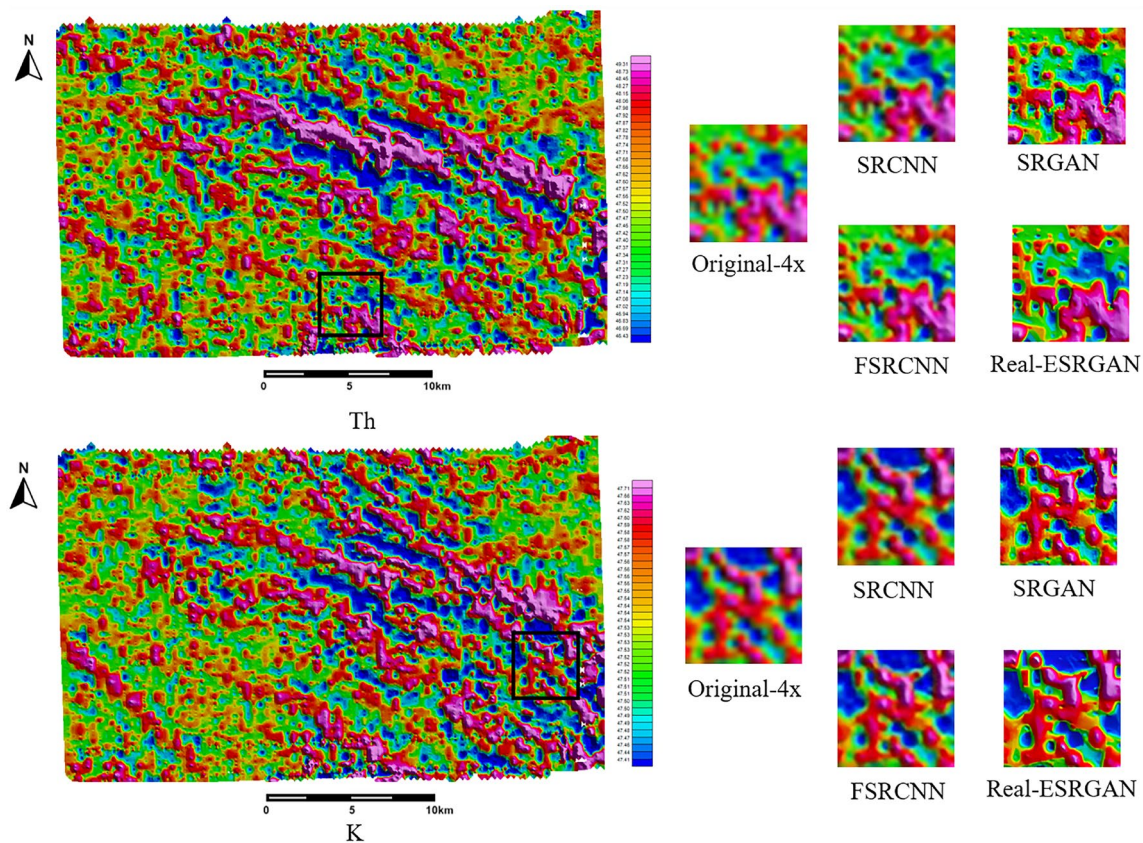


Fig. 6 (Color online) Comparison of results of 4× super-resolution reconstruction without algorithms

distinctly visible, with a significant reduction in the sense of image blurriness. In stark contrast to SRCNN, FSRCNN, and SRGAN, the Real-ESRGAN algorithm demonstrates a pronounced superiority in reconstructed image fidelity, efficaciously attenuating artifacts such as ringing and overshoot. As portrayed in Fig. 5, the algorithm enables the distinct identification of depositional loci across macroscopic and microscopic scales. The reconstructed imagery not only elucidates the ore body's continuity but also accentuates the demarcation between the ore body and the enveloping rock matrix, affording an exact visual portrayal of the ore deposit's morphology, magnitude, and orientation. However, an increase in magnification intensity introduces an oversaturation at the ore body periphery, as noted in Fig. 7.

In order to comprehensively verify the superiority of Real-ESRGAN algorithm, this study further selected and reconstructed U, Th, K and Tc images in the test set, as shown in Fig. 8. The results show that the image reconstructed by the Real-ESRGAN algorithm has a significant improvement in overall sharpness and color vividness. A smoother and more natural transition is displayed at the edge of the ore body and key geological landmarks, effectively avoiding the misunderstanding and error that may arise in the image processing. This not only improves the overall

resolution of the image, but more importantly, enhances the detailed representation of geological structures and deposit sites, providing more abundant and accurate identifiable features for geological exploration and resource evaluation.

When evaluating the quality of super-resolution reconstructed images, in addition to the subjective analysis based on visual senses, this study also adopted a variety of objective quantitative indicators to verify the accuracy of experimental results. Not only the peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) are introduced as evaluation criteria, but the error between the reconstructed image and the original image is objectively quantified, as shown in Tables 1 and 2. Residual analysis between the original image and the generated super-resolution image has also been added by calculating the residual image between the two and analyzing its histogram, the ability of Real-ESRGAN algorithm in detail recovery and noise processing is revealed, which provides a comprehensive perspective for evaluating the quality of super-resolution reconstructed images as shown in Fig. 9.

The analysis results from Tables 1 and 2 show that with the increase in magnification, the PSNR and SSIM values of reconstructed images obtained by each algorithm generally show a downward trend. Notably, the Real-ESRGAN

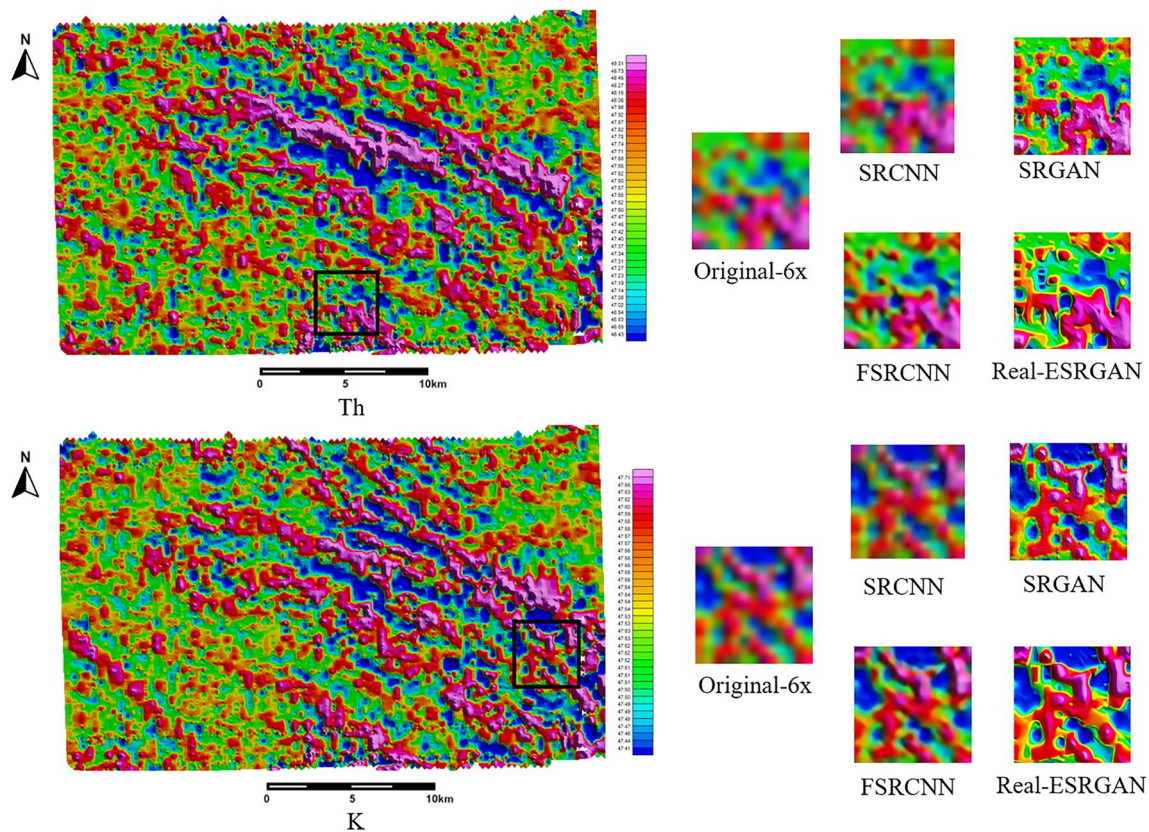


Fig. 7 (Color online) Comparison of results of 6× super-resolution reconstruction without algorithms

algorithm demonstrates a pronounced superiority in both indices, signifying its exceptional performance in reconstructing super-resolution images. When juxtaposing these quantitative measures, it is evident that the Real-ESRGAN algorithm's reconstructions are markedly closer to the original images, particularly observed in the SSIM values, which excel in capturing the structural fidelity and perceptual quality of the imagery. The SSIM metric, with its sensitivity to both local alterations and the granularity of textures, provides a comprehensive assessment of image quality. Illustratively, at a 2× magnification, the Real-ESRGAN algorithm achieves an SSIM value of 0.950, a figure that underscores the close resemblance of the reconstructed image to its pristine counterpart. This high SSIM value not only attests to the Real-ESRGAN algorithm's proficiency in preserving fine details but also substantiates its overall efficacy and preeminence in the realm of super-resolution image reconstruction tasks.

It can be seen from Fig. 9 that the residual values are mainly concentrated in the central position, and most of the residual values are close to zero. This shows that the difference between the super-resolution reconstructed image and the original image is very small on most pixels, and the Real-ESRGAN model can better retain the distribution

characteristics of radioactive elements in the original image. Although most of the residuals are concentrated in the center, it can also be seen that some residuals are too large, which may be due to the differences between the reconstructed image and the original image of the geological boundary, high-frequency structure or noisy area. The symmetry of the histogram reflects that the residuals are evenly distributed in positive and negative directions. This symmetry means that the Real-ESRGAN model has no obvious systematic bias and can restore the radiation intensity change of the UAV-borne gamma-ray spectrum image in a balanced way.

5 Conclusion

In this study, the Real-ESRGAN algorithm is successfully applied to reconstruct the UAV-borne gamma-ray spectrum image in super-resolution, significantly enhancing the spatial resolution of the image and improving the visualization quality of geological features. By comparing with SRCNN, SRGAN and FSRCNN algorithms, Real-ESRGAN has shown excellent performance in PSNR and SSIM objective evaluation indices. Especially, under the condition of

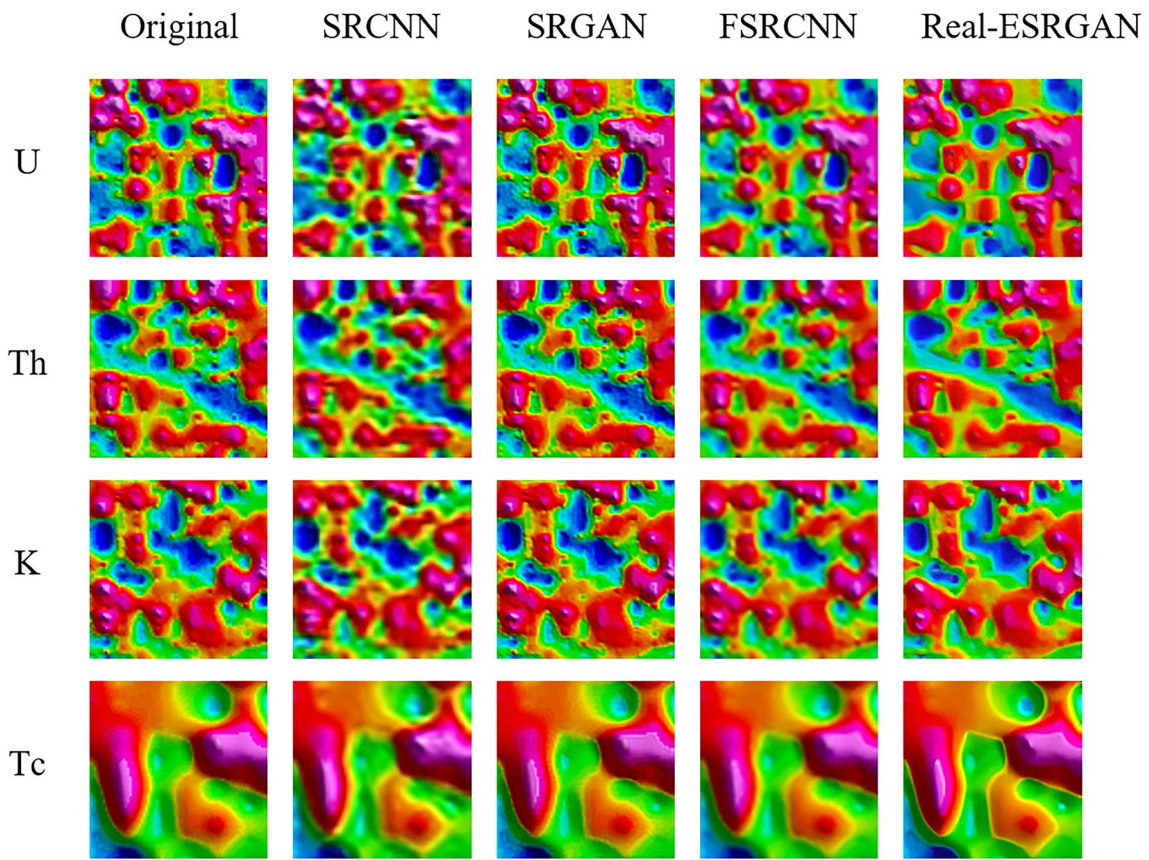


Fig. 8 (Color online) Comparison of 2× super-resolution image reconstruction results with different algorithms

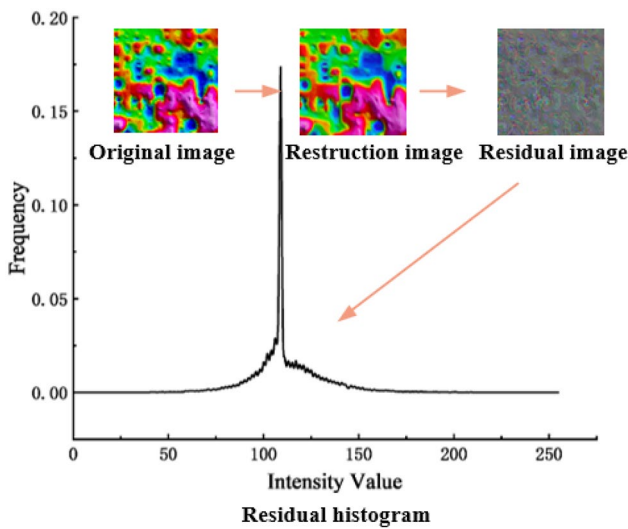


Fig. 9 (Color online) Residual difference between original image and super-resolution reconstructed image

Table 1 Comparison of PSNR of super-resolution reconstructed images by different algorithms

Magnification	SRCNN	SRGAN	FSRCNN	Real-ESRGAN
2	16.914	26.176	20.215	29.176
4	15.167	25.109	16.942	26.408
6	12.907	24.326	15.177	23.827

Table 2 Comparison of SSIM of super-resolution reconstructed images by different algorithms

Magnification	SRCNN	SRGAN	FSRCNN	Real-ESRGAN
2	0.659	0.934	0.802	0.950
4	0.502	0.906	0.638	0.942
6	0.272	0.868	0.502	0.938

2× amplification, the SSIM value is as high as 0.950, which substantiates its advantage in detail preservation and texture clarity, and further highlights the significant enhancement in

the identification of geological body boundaries. Additionally, the Real-ESRGAN algorithm effectively reduces ringing and overshoot artifacts, making the transition between

the edge of the ore body and key geological markers more smooth and natural and greatly enhancing the detailed representation of the geological body structure and deposit sites. This clear delineation of lithological boundaries provides geologists with more intuitive and accurate geological information, thereby offering significant application value in geological exploration and resource assessment. Consequently, the Real-ESRGAN algorithm is not only theoretically advanced but also demonstrates robust practical utility, providing an effective image processing tool for UAV-borne gamma-ray spectrum image processing.

Author Contributions All authors contributed to the study conception and design. Data collection and analysis were performed by Xin Wang, Guang-Hao Luo, and Qi-Qiao Wei. Early-stage literature investigation was conducted by Yuan Yuan and Xuan Zhao. Data curation and experimental guidance were provided by He-Xi Wu and Chao Xiong. The first draft of the manuscript was written by Xin Wang, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data Availability The data that support the findings of this study are openly available in Science Data Bank at <https://cstr.cn/31253.11.sciencedb.j00186.00836> and <https://www.doi.org/10.57760/sciencedb.j00186.00836>.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

- H.M. Ba, M. Jaffal, K. Lo et al., Mapping mafic dyke swarms, structural features, and hydrothermal alteration zones in Atar, Ahmeyim and Chami areas (Reguibat Shield, Northern Mauritania) using high-resolution aeromagnetic and gamma-ray spectrometry data. *J. Afr. Earth Sci.* **163**, 103749 (2020). <https://doi.org/10.1016/j.jafrearsci.2019.103749>
- S.M. Hassan, M.A.S. Youssef, S.S. Gabr et al., Radioactive mineralization detection using Remote Sens-Basel and airborne gamma-ray spectrometry at Wadi Al-Miyah area, Central Eastern Desert, Egypt. *Egypt. J. Remote Sens-Basel Space Sci.* **25**, 37–53 (2022). <https://doi.org/10.1016/j.ejrs.2021.12.004>
- M. Ohera, L. Gryc, I. Češpírová et al., Airborne HPGe spectrometer for monitoring of air dose rates and surface activities. *Nucl. Eng. Technol.* **55**, 4039–4047 (2023). <https://doi.org/10.1016/j.net.2023.07.019>
- S.H. Zhang, Z.Y. Lin, P. Soc, Photo-opt ins and analysis of airborne spectral data in Xiangshan uranium ore field. *J. East China Univ. Technol. Nat. Sci. Ed.* **35**(02), 124–128 (2012) <https://doi.org/CNKI: SUN: HDDZ.0.2012-02-005>
- O. Salek, M. Matolin, L. Gryc, Mapping of radiation anomalies using UAV mini-airborne gamma-ray spectrometry. *J. Environ. Radioact.* **182**, 101–107 (2018). <https://doi.org/10.1016/j.jenvrad.2017.11.033>
- B.H. Li, Y.Y. Lu, G.Y. Zhang et al., Research on the application of curvelet transform in airborne gamma spectral data process-London. *World Nucl. Geosci.* **40**(4), 963–973 (2023). <https://doi.org/10.3969/j.issn.1672-0636.2023.04.007>
- Y. Lyu, W. Wang, P. Chen, Fixed-Wing UAV based air-to-ground channel measurement and modeling at 2.7 GHz in rural environment. *IEEE T. Antenn. Propag.* **73**, 2038–2052 (2025). <https://doi.org/10.1109/TAP.2024.3428337>
- J. Wang, Q. Zhu, Z. Lin et al., Sparse Bayesian learning-based hierarchical construction for 3D radio environment maps incorporating channel shadowing. *IEEE Trans. Wirel. Commun.* **23**, 14560–14574 (2024). <https://doi.org/10.1109/TWC.2024.3416447>
- Y. Xiao, Q. Yuan, K. Jiang et al., Frequency-assisted mamba for remote sens-basel image super-resolution. (2024). <https://doi.org/10.48550/arXiv.2405.04964>
- F. Nan, W. Jing, F. Tian et al., Feature super-resolution based facial expression recognition for multi-scale low-resolution images. *Knowl.-Based Syst.* **236**, 107678 (2022). <https://doi.org/10.1016/j.knsys.2021.107678>
- Y. Xu, X. Li, Y. Jie et al., Simultaneous tri-modal medical image fusion and super-resolution using conditional diffusion model (2024). <https://doi.org/10.48550/arXiv.2404.17357>
- C. Dong, C.C. Loy, K. He et al., Learning a deep convolutional network for image super-resolution, in *Computer Vision—ECCV 2014*. ECCV 2014. Lecture Notes in Computer Science, vol. 8692 (Springer, Cham, 2014), pp.184–199. https://doi.org/10.1007/978-3-319-10593-2_13
- W.Z. Shi, J. Caballero, F. Huszár et al., Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2016), pp. 1874–1883. <https://doi.org/10.1109/CVPR.2016.207>
- J. Kim, J.K. Lee, K.M. Lee, Accurate image super-resolution using very deep convolutional networks, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2016), pp. 1646–1654. <https://doi.org/10.1109/CVPR.2016.182>
- C. Ledig, L. Theis, F. Huszár et al., Photo-realistic single image super-resolution using a generative adversarial network, in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2017), pp. 105–114. <https://doi.org/10.1109/CVPR.2017.19>
- J. Zhang, J.Y. Li, X.Y. Li et al., Image super-resolution reconstruction based on SRGAN. *Comput. Knowl. Technol.* **20**(01), 14–17 (2024). <https://doi.org/10.14004/j.cnki.ckt.2024.0029>
- Z.W. Zhou, Research on the super-resolution of engineering cost bills based on SRGAN. *Intell. City* **9**(10), 105–107 (2023). <https://doi.org/10.19301/j.cnki.zncs.2023.10.033>
- K. Zhang, W. Zuo, L. Zhang, Learning a single convolutional super-resolution network for multiple degradations, in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (IEEE, 2018), pp. 3262–3271. <https://doi.org/10.1109/CVPR.2018.00344>
- X.C. Hu, Research on adaptive image super-resolution algorithms. Dissertation, University of Science Technology of China, 2021. <https://doi.org/10.27517/d.cnki.gzkju.2021.000009>
- G.L. Wu, Research on blind image super-resolution algorithms based on deep learning. Dissertation, Xiamen University, 2021. <https://doi.org/10.27424/d.cnki.gxmd.2021.001765>
- K. Zhang, J. Liang, L.V. Gool et al., Designing a practical degradation model for deep blind image super-resolution (2021). <https://doi.org/10.48550/arXiv.2103.1S4006>
- Z. Luo, H. Huang, L. Yu et al., Deep constrained least squares for blind image super-resolution (2022). <https://doi.org/10.48550/arXiv.2202.07508>
- Y. Yuan, S. Liu, J. Zhang et al., Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks (IEEE, 2018). <https://doi.org/10.1109/CVPRSW.2018.00113>
- Y.F. Niu, D.Y. Wang, P. Mostaghimi et al., An innovative application of generative adversarial networks for physically accurate

- rock images with an unprecedented field of View. *Geophys. Res. Lett.* **47**(23), e2020GL089029 (2020). <https://doi.org/10.1029/2020GL089029>
25. A. Liu, Y. Liu, J. Gu et al., Blind image super-resolution: a survey and beyond. *IEEE Trans. Pattern Anal.* **45**, 5461–5480 (2022). <https://doi.org/10.1109/TPAMI.2022.3203009>
 26. A. Bulat, J. Yang, G. Tzimiropoulos, To learn image super-resolution, use a GAN to learn how to do image degradation first, in *European Conference on Computer Vision*. (Springer, Cham, 2018). https://doi.org/10.1007/978-3-030-01231-1_12
 27. M. Fritsche, S. Gu, R. Timofte, Frequency separation for real-world super-resolution, in *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)* (IEEE, 2019), pp. 3599–3608. <https://doi.org/10.1109/ICCVW.2019.00445>
 28. Y. Zhou, W. Deng, T. Tong et al., Guided frequency separation network for real-world super-resolution, in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (IEEE, 2020), pp. 1722–1731. <https://doi.org/10.1109/CVPRW50498.2020.00222>
 29. X. Wang, L. Xie, C. Dong et al., Real-esrgan: training real-world blind super-resolution with pure synthetic data (2021). <https://doi.org/10.48550/arXiv.2107.10833>
 30. X. Wang, K. Yu, S. Wu et al., Esrgan: enhanced super-resolution generative adversarial networks (2018). https://doi.org/10.1007/978-3-030-11021-5_5
 31. J.P.F. D’Haeyer, Gaussian filtering of images: a regularization approach. *Signal Process.* **18**(2), 169–181 (1989). [https://doi.org/10.1016/0165-1684\(89\)90048-0](https://doi.org/10.1016/0165-1684(89)90048-0)
 32. M. Kirchner, J. Fridrich, On detection of median filtering in digital images. *Int. Soc. Opt. Photon.* **7541**, 754110 (2010). <https://doi.org/10.1117/12.839100>
 33. A. Ruangma, B. Bai, J.S. Lewis et al., Three-dimensional maximum a posteriori (MAP) imaging with radiopharmaceuticals labeled with three Cu radionuclides. *Nucl. Med. Biol.* **33**(2), 217–226 (2006). <https://doi.org/10.1016/j.nucmedbio.2005.11.001>
 34. M. Jérémie, C. Phillips, W.D. Penny et al., MEG source localization under multiple constraints: an extended Bayesian framework. *Neuroimage* **30**(3), 753–767 (2007). <https://doi.org/10.1016/j.neuroimage.2005.10.037>
 35. Y. Kataoka, T. Matsubara, K. Uehara, Image generation using generative adversarial networks and attention mechanism, in *2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)* (IEEE, 2016), pp. 1–6. <https://doi.org/10.1109/ICIS.2016.7550880>
 36. X. Zhang, C. Feng, A. Wang et al., CT super-resolution using multiple dense residual block based GAN. *SIViP* **15**, 725–733 (2021). <https://doi.org/10.1007/s11760-020-01790-5>

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.