Super-Resolution Reconstruction of UAV-borne Gamma-Ray Spectrum Images Based on Real-ESRGAN Algorithm

Authors: Wang Xin, Yuan Yuan, Zhao Xuan, Luo Guanghao, Wei Qiqiao, Wu Hexi, Xiong Chao, Xiong Chao

Date: 2025-07-28T11:30:03+00:00

Abstract

Unmanned aerial vehicle (UAV)-borne gamma-ray spectrum surveys play 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 interpretation. 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 $2\times$, $4\times$, and $6\times$ 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 $2\times$ magnification, significantly higher than that of 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.

Full Text

Preamble

Super-resolution reconstruction of UAV-borne Gamma-Ray spectrum images based on Real-ESRGAN algorithm

Xin Wang,[1] Yuan Yuan,[1] Xuan Zhao,[1] Guang-Hao Luo,[1] Qi-Qiao Wei,[1] He-Xi Wu,[1] and Chao Xiong[1],†

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

Unmanned aerial vehicle (UAV)-borne gamma-ray spectrum surveys play 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 interpretation. This study explores the application of the Real-ESRGAN algorithm in the super-resolution reconstruction of UAV-borne gamma-ray spectrum images to enhance spatial resolution and improve the quality of geological feature visualization. We conducted super-resolution reconstruction experiments with $2\times$, $4\times$, and $6\times$ magnification using the Real-ESRGAN algorithm, comparing the results with three other mainstream algorithms (SRCNN, SRGAN, FSRCNN) to verify its superiority in image quality. The experimental results indicate that Real-ESRGAN achieved a structural similarity index (SSIM) value of 0.950 at $2\times$ magnification, significantly higher than that of 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

Introduction

UAV-borne gamma-ray spectrum surveys measure and record the type and intensity of gamma rays emitted by natural radionuclides (e.g., uranium, thorium, and potassium) from the ground and shallow subsurface. These data, combined with coordinate information, are processed, screened, and corrected to produce maps of surface radionuclide distributions, which are essential for geological exploration, radioactive mineral prospecting, and environmental monitoring. Compared with fixed-wing aircraft surveys, UAV-borne gamma-ray spectrum surveys offer lower flight costs, higher measurement efficiency, safer and more flexible flight operations, and prolonged hovering capabilities. Consequently, they have been widely adopted in radioactive mineral exploration, radiation environment monitoring, nuclear emergency response, and related fields [?, ?, ?].

Zhang Shihong [?] skillfully employed the tension spline method and Munsell transform technology to perform fine rasterization and anomaly correction of aerial gamma spectrum data, successfully extracting key information and clearly displaying the energy spectrum characteristics of the Xiangshan volcanic basin. By analyzing the distribution of uranium, thorium, and potassium, this work provided new clues for uranium exploration and advanced the development of geological exploration technology. Ondřej Salek et al. [?] tested the performance of a new type of small airborne gamma-ray spectrum measurement equipment

based on UAVs and verified its ability to detect uranium anomalies at different altitudes. They found that it accurately captured small changes in gamma ray intensity even at high altitudes. Li et al. [?] developed a curvelet-based noise reduction technique for airborne gamma data that can reconstruct multiscale data for different analysis purposes, effectively removing noise while retaining local anomaly information to avoid resolution loss and boundary effects.

Lyu et al. [?] proposed a K-factor prediction model that considers flight altitude and direct distance, paying special attention to key factors such as path loss, shadow fading, and small-scale fading at different flight altitudes, which could potentially be applied to aerial gamma-ray spectral image processing. Wang et al. [?] proposed a layered approach specifically for the radio environment map (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 serve as references in research on aerial gamma-ray spectral image processing to improve processing accuracy and efficiency.

The demand for high-resolution UAV-borne gamma-ray spectrum images has increased to improve research accuracy. 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, 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 detail and information content of existing images. This technology plays an important role in many fields such as satellite imagery [?], face recognition [?], and medical imaging [?]. In recent years, with the development of deep learning, it has been used to improve image super resolution. SR methods based on convolutional neural networks (CNNs), residual networks (ResNets), and generative adversarial networks (GANs) have been proposed.

SRCNN [?], the first SR deep learning network, achieved fast online applications owing to its lightweight structure and excellent recovery quality, despite its large computational load. ESPCN [?] introduced an innovative subpixel convolution layer to obtain HR from LR at minimal computational cost. VDSR [?] uses an extremely deep convolutional network combined with residual learning to significantly improve convergence speed during training. These methods improve the accuracy and speed of image SR by using faster and deeper CNNs. However, when enlarging an LR image to an HR image, the reconstructed SR image often lacks texture details due to large-scale factors, resulting in unsatisfactory reconstruction effects. SRGAN [?, ?, ?] enhanced the content loss function with adversarial loss by training a GAN and replaced the content loss based on mean squared error (MSE) with loss based on a VGG network feature map, effectively overcoming the problem of low perceptual quality of reconstructed images and

making the generated images more realistic. However, this method suffers from complex network structure and lengthy training process.

The degradation of image quality is a complex phenomenon, usually due to limitations of the imaging system and environmental disturbances. Blind super-resolution (Blind SR) technology can effectively restore LR images to HR images when the degradation process is unclear. This technology can be divided into two categories: explicit and implicit modeling. Explicit modeling parameterizes blur kernel and noise information. SRMD [**?**, **?**, **?**], the first deep-learning-driven Blind SR method, introduces a dimension expansion strategy that enables the convolution network to handle blur kernel and noise level inputs and solves the dimension mismatch problem, although it may perform poorly for uncovered degradation types. The BSRGAN model [**?**], proposed by Zhang et al. in 2021, enhances model adaptability to real-world degradation by introducing complex degradation factors and random shuffling strategies. Based on the kernel stretching strategy of Zhang et al., Luo et al. [**?**] proposed a new practical degradation model that uses a dynamic depth linear filter and a constrained least squares deconvolution algorithm based on a neural network to improve the restoration quality of blurred images. Implicit modeling abandons explicit parameters and uses deep learning techniques, particularly GANs, to restore LR images directly to HR images. The CinCGAN [**?**, **?**] model adopts a double-cycle GAN structure that effectively solves the problem of complex interference in LR inputs. The research classification of Liu et al. [**?**] 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 [**?**], FSSR [**?**], and FSSR-GAN [**?**]. However, these methods may have domain gap problems. DASR improves SR training performance through domain-gap awareness training and a domain-distance-weighted monitoring strategy. Real-ESRGAN [**?**] improved upon ESRGAN [**?**] using a discriminator designed with U-Net and a high-order degradation process, introducing a sinc filter to reduce ringing and overshoot artifacts. It provides a more accurate and stable SR 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 HR images. Real-ESRGAN technology can improve the clarity and feature enhancement of UAV-borne gamma-ray spectrum images and help interpret geological data more accurately. For example, it can enhance areas of an image that are blurred due to topographic effects, making geological features more clearly visible. In addition, the denoising and detail enhancement capabilities of Real-ESRGAN help remove the noise generated during flight measurements, thereby improving data quality. Image SR technology can improve LR images to HR images through algorithmic processing and compensate for hardware equipment shortages. This technology, combined with UAV-borne gamma-ray spectrum survey, not only improves the clarity and detail expression of the image but also effectively improves the accuracy and reliability of data analysis. Therefore, the combination of UAV-borne gamma-ray

spectrum survey and SR image technologies can more effectively monitor environmental radiation levels, provide more accurate data support for the efficient exploration of radioactive minerals, and promote the development of radioactive geophysical exploration technology.

Technical Principles and Methods

UAV-borne Gamma-ray Spectrum Technology

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 environmental radiation monitoring. However, the generated images 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 typically required to improve the quality of UAV-borne gamma-ray spectrum images. A Gaussian filter, which reduces the interference of noise on image quality by smoothing the image, is widely used. The formula for Gaussian filtering [**?**] can be expressed as:

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

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 the pixels in the image, and σ denotes the standard deviation of the Gaussian distribution. However, when the Gaussian filter is used for global smoothing of the entire image, it lacks self-adaptation and cannot smooth different regions according to local features.

Median filtering [**?**] is also a common nonlinear filtering technique that 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:

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

where $I(x, y)$ denotes the pixel value of the original image at position (x, y), $I'(x, y)$ is the filtered image pixel value at position (x, y), and W is a neighborhood window, typically a $(2k + 1) \times (2k + 1)$ rectangular window centered at (x, y). The median represents the median of the pixel values 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 the geological features of 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 an image, which is critical for determining the precise locations of geological structures. However, these techniques may increase artificial distortion of images, leading to overemphasis or misidentification of geological features.

Image Super-resolution Reconstruction

Image SR aims to convert LR images into HR images by modeling the degradation process, which includes blurring, down-sampling, and noise. The goal is to determine 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 [**?**], which involves a least-squares problem with a regularization term to incorporate prior knowledge. The core goal of image SR reconstruction technology is to convert LR images into HR images. Mathematically, this process is often modeled as a degradation process in which the LR image is regarded as the result of the HR image being blurred, down-sampled, and disturbed by noise:

$$I_{LR} = H \cdot I_{HR} + n$$

where H is a degenerate operator that includes processes such as downsampling, blurring, and noise, and can be expressed as a matrix or an integral operator. n represents the observation noise, which may include instrument noise and environmental noise. The general production is assumed to follow a Gaussian or Poisson distribution.

The goal of SR reconstruction is to find an operator R that makes $R(I_{LR})$ as close as possible to the original HR image I_{HR}. The formula is as follows:

$$\hat{I}_{HR} = R(I_{LR})$$

where $\hat{I}_{HR}$ denotes the estimated HR image.

Under the Bayesian framework [**?**], SR reconstruction can be expressed as a MAP problem:

$$R^* = \arg\max p(I_{HR}|I_{LR})$$

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

In practice, because it is usually not feasible to compute $p(I_{HR}|I_{LR})$ directly, the SR reconstruction algorithm typically uses a regularization term that translates into a least-squares problem:

$$R^* = \arg\min \left\{ \|I_{LR} - H \cdot R(I_{LR})\|_2^2 + \lambda \cdot \Omega(R(I_{HR})) \right\}$$

where $\| \cdot \|_2$ denotes the L_2 norm and is used to measure the reconstruction error, $\Omega(\cdot)$ is a regularization term (such as gradient smoothing or sparse representation) used to introduce prior knowledge of I_{HR}, and 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 deep neural network. Simultaneously, a GAN is used to improve the visual quality and naturalness of the reconstructed image.

Super-resolution Reconstruction of UAV Gamma-ray Spectrum Images

Real-ESRGAN Principle

Enhanced super-resolution generative adversarial networks (ESRGANs), which are based on GANs, generate high-quality SR images through competition between the generator and discriminator. Real-ESRGAN improves upon ESRGAN specifically for the SR 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. 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.

Generative Adversarial Network (GAN) The GAN [**?**] consists of two main components: generator and discriminator. As shown in [Figure 1: see original paper], the generator converts the LR image into an HR image, whereas the discriminator distinguishes between the generated and real HR images. 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 capability to distinguish between generated and real images more effectively.

Network Architecture Real-ESRGAN retains the residual-in-residual dense block (RRDB) [**?**] from ESRGAN as the core component of its generator. Through an innovative network structure design, the performance of image SR reconstruction is significantly improved. The RRDB is composed of multiple residual blocks, each of which is further densely connected, allowing features

from all previous layers to be directly connected to the current layer. The 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 improved detail clarity in the generated images. In addition, the RRDB' s role in the 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 layers and feature fusion modules, to generate high-quality HR images. [Figure 2: see original paper] shows the structure of the RRDB. Through this design, Real-ESRGAN is able to produce detailed and visually realistic SR images to meet the needs of various practical applications.

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 (MAE)

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

$$\text{L1 Loss} = \sum \left| I_{HR}^{(i)} - \hat{I}_{HR}^{(i)} \right|$$

where $I_{HR}^{(i)}$ is the ith real HR image pixel value, $\hat{I}_{HR}^{(i)}$ corresponds to the generated image pixel value, and N is the total number of pixels.

2) Perceptual Loss

Perceptual loss is typically based on pretrained CNNs (such as VGG networks) to extract features and compare the differences in these features between generated and real images. A simplified perceptual loss can be expressed as:

$$\text{Perceptual Loss} = \sum \left\| F\left(I_{HR}^{(i)}\right) - F\left(\hat{I}_{HR}^{(i)}\right) \right\|_2^2$$

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

3) Adversarial Loss

By training the discriminator to distinguish between real and generated images, the generator' s goal is to maximize the probability that the discriminator will make an incorrect judgment. A common form of adversarial loss uses Wasserstein distance, which is expressed as:

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

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

These loss functions work synergistically to ensure that the generated HR images not only exhibit fidelity at the pixel level but are also visually realistic and sufficiently detailed to meet the needs of various real-world applications.

Data Source

The data used in this study were obtained from an experimental area in northern Gansu Province, China, as shown in [Figure 3: see original paper]. In total, 62,889 data points from the total count (Tc), uranium (U), thorium (Th), and potassium (K) channels were measured and processed to create equivalent maps of surface radionuclide distributions. To enhance the visual quality and detailed resolution of the UAV-borne gamma-ray spectrum images, image restoration techniques were employed on the raw data. 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 to measure radioactive count rates quantitatively.

In our study, to adapt to the specific resolution and characteristics of 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 exceeded the image boundaries, ultimately forming 232 two-dimensional slices of 80×80 pixels each. The 80×80 pixel cutting size was selected to take full advantage of the spatial resolution of the gamma-ray detector at the actual flight altitude while maintaining sufficient detail for effective geological feature analysis.

To obtain LR 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 resampling algorithm, which is a high-quality resampling method suitable for image scaling. As shown in [Figure 4: see original paper], with a six-fold increase in the UAV-borne gamma spectrum image, some geological detail loss can be clearly observed.

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 Real-ESRGAN model was constructed using the PyTorch framework. The Adam optimizer was used to train the model at a learning rate of 1×10^{-4}, and exponential moving average (EMA) was employed for more stable training. In

addition, L1 loss, perceptual 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 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 improves the effects of UAV-borne gamma-ray spectrum SR image reconstruction.

UAV-borne gamma-ray spectrum images are a special type of remote sensing image that provide 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 require high spatial resolution to capture subtle geological features. The main advantage of SR images is their ability to significantly enhance the spatial resolution of the image, making the visual effect clearer, more detailed, and more realistic. To evaluate the application of SR reconstruction technology to UAV-borne gamma-ray spectrum images, this study selected the peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) as the main evaluation indicators.

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

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

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 for PSNR is:

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

where MAX denotes 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 values corresponding to the images in this study were 255.

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

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

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 grayscales of the original image x and the reconstructed image y, and C_1 is a non-zero constant; $c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$ 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, and C_2 is a non-zero constant; $s(x,y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$ is structural comparison, σ_{xy} is the covariance of the grayscale of the original image x and the reconstructed image y, and C_3 is usually $C_2/2$. α, β, and γ are weight parameters, usually taking the value of 1. The final SSIM index formula is as follows:

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

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. In contrast, the lower the SSIM value, the greater the difference between the reconstructed and original images, and the worse the quality.

Results and Discussion

In this study, we reconstructed UAV-borne gamma-ray spectrum images with 2×, 4×, and 6× SR. The SRCNN, SRGAN, FSRCNN, and Real-ESRGAN algorithms were compared to verify the effectiveness of the Real-ESRGAN algorithm.

To validate model performance, 10 images were randomly selected from the image set of four elements (U, Th, K, and Tc) to constitute a test set. Th and K were selected for analysis during the testing process. By comparing and analyzing the geological map of the survey area (as shown in [Figure 3: see original paper]), the distribution patterns of Th and K within the survey area exhibited a high degree of consistency. Notably, a fault within the survey area extends from northwest to southeast, which aligns closely with the high-value bands of Th and K. This phenomenon is attributed to the distribution of Hercynian late-stage granites, which caused anomalous enrichment of radioactive K and Th in the survey area. Typically, granites contain minerals such as potassium feldspar and mica, which are rich in K; concurrently, they also contain Th-bearing minerals such as monazite and xenotime. Consequently, in areas where granite is present, the contents of thorium and potassium are generally high. Therefore, the test dataset employed in this study accurately reflected the distribution of natural radioactive nuclides within the survey area.

To visually demonstrate the superior effect of the Real-ESRGAN algorithm on SR reconstruction of UAV-borne gamma-ray spectrum images, the three

introduced SR reconstruction algorithms were applied with 2×, 4×, and 6× magnification to the test set. The results are presented in [Figure 5: see original paper], [Figure 6: see original paper], and [Figure 7: see original paper].

The analyses presented in these figures expose discernible disparities in the efficacy of the assorted algorithms engaged in imagery reconstruction. The SR reconstructions yielded by the SRCNN algorithm exhibited a measure of obscurity and chromatic aberration with pronounced exacerbation in blurriness concomitant with escalated magnification. In particular, at the 6× magnification exemplified in [Figure 7: see original paper], the loss of detailed 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 refinement of convolutional strata and activation mechanisms, thereby resulting in marked enhancement in the textural, marginal, and configurational fidelity of the reconstructed imagery relative to those of SRCNN (as depicted in [Figure 5: see original paper]). Despite these advancements, discrepancies persist between FSRCNN's output and the authentic geological background and boundaries, with observable residual haziness (as illustrated in [Figure 7: see original paper]).

The SRGAN algorithm surpasses its precursors in holistic quality of image reconstruction by integrating the GAN architecture to facilitate more nuanced rendition of intricate image details. Even with magnification increased to 6×, the detailed information and boundaries of geological bodies remain distinctly visible, with significant reduction in image blurriness. In stark contrast to SRCNN, FSRCNN, and SRGAN, the Real-ESRGAN algorithm demonstrates pronounced superiority in reconstructed image fidelity, efficaciously attenuating artifacts such as ringing and overshoot. As portrayed in [Figure 5: see original paper], the algorithm enables 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 exact visual portrayal of the ore deposit's morphology, magnitude, and orientation. However, an increase in magnification intensity introduces over-saturation at the orebody periphery, as shown in [Figure 7: see original paper].

To comprehensively verify the superiority of the Real-ESRGAN algorithm, this study further selected and reconstructed U, Th, K, and Tc images in the test set, as shown in [Figure 8: see original paper]. The results indicate that images reconstructed by the Real-ESRGAN algorithm show 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 misunderstandings and errors that may arise during image processing. This not only improves overall image resolution but also enhances detailed representation of geological structures and deposit sites, providing more abundant and accurate identifiable features for geological exploration and resource evalu-

ation.

When evaluating the quality of super-resolution reconstructed images, in addition to subjective analysis based on visual senses, this study also adopted various objective quantitative indicators to verify experimental accuracy. PSNR and SSIM were introduced as evaluation criteria to objectively quantify the error between reconstructed and original images, as presented in and . A residual analysis between the original and generated SR image was also conducted by calculating the residual image and analyzing its histogram. The ability of the Real-ESRGAN algorithm in detail recovery and noise processing was revealed, providing a comprehensive perspective for evaluating SR reconstructed image quality, as shown in [Figure 9: see original paper].

Comparison of the peak signal-to-noise ratio (PSNR) of super-resolution reconstructed images using different algorithms

Magnification	SRCNN	SRGAN	FSRCNN	Real-ESRGAN

Comparison of the structural similarity index (SSIM) of super-resolution reconstructed images using different algorithms

Magnification	SRCNN	SRGAN	FSRCNN	Real-ESRGAN

The analysis results from and indicate that with an increase in magnification, the PSNR and SSIM values of reconstructed images obtained by each algorithm generally show a downward trend. Notably, the Real-ESRGAN algorithm demonstrates pronounced superiority in both indices, signifying its exceptional performance in reconstructing SR images. When juxtaposing these quantitative measures, it is evident that Real-ESRGAN algorithm reconstructions are markedly closer to original images. This is particularly observed in SSIM values, which excel in capturing structural fidelity and perceptual quality. The SSIM metric, sensitive to both local alterations and texture granularity, provides comprehensive assessment of image quality.

Illustratively, at $2\times$ magnification, the Real-ESRGAN algorithm achieves an SSIM value of 0.950, which underscores the close resemblance of the reconstructed image to its pristine counterpart. This high SSIM value not only attests to Real-ESRGAN' s proficiency in preserving fine details but also substantiates its overall efficacy and preeminence in SR image reconstruction tasks.

As observed from [Figure 9: see original paper], residual values are mainly concentrated in the central position and most are close to zero. This shows that the difference between the SR reconstructed image and original image is very small for most pixels, and the Real-ESRGAN model can better retain the distribution characteristics of radioactive elements in the original image. Although

most residuals are concentrated in the center, some residuals are too large, which may be due to differences between reconstructed and original images at geological boundaries, high-frequency structures, or noisy areas. The symmetry of the histogram indicates that 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 radiation intensity changes in UAV-borne gamma-ray spectrum images in a balanced manner.

Conclusion

In this study, the Real-ESRGAN algorithm was successfully applied to reconstruct UAV-borne gamma-ray spectrum images at super-resolution, significantly enhancing spatial resolution and improving visualization quality of geological features. By comparing it with SRCNN, SRGAN, and FSRCNN algorithms, Real-ESRGAN demonstrated excellent performance in PSNR and SSIM objective evaluation indices. In particular, under $2\times$ magnification, the SSIM value was as high as 0.950, substantiating its advantage in detail preservation and texture clarity and further highlighting significant enhancement in identification of geological body boundaries.

Additionally, the Real-ESRGAN algorithm effectively reduced ringing and overshoot artifacts, making transitions between ore body edges and key geological markers smoother and more natural, and greatly enhancing detailed representation of geological body structures and deposit sites. This clear delineation of lithological boundaries provides geologists with more intuitive and accurate geological information, thereby offering significant application value for 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.

References

[1] 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). doi: 10.1016/j.jafrearsci.2019.103749

[2] 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. The Egyptian Journal of Remote Sens-Basel and Space Sciences. 25, 37–53 (2022). doi: 10.1016/j.ejrs.2021.12.004

[3] 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). doi: 10.1016/j.net.2023.07.019

[4] S.H. Zhang and Z.Y. Lin, P. Soc Photo-Opt Ins and analysis of airborne spectral data in Xiangshan uranium ore field. Journal of East China University of Technology: Natural Science Edition 35(02), 124–128 (2012). doi: CNKI: SUN: HDDZ.0.2012-

[5] O. Salek, M. Matolin and L. Gryc, Mapping of radiation anomalies using UAV mini-airborne gamma-ray spectrometry. J. Environ. Radioact. 182, 101–107 (2018). doi: 10.1016/j.jenvrad.2017.11.033

[6] 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 Nuclear Geoscience 40(4), 963–973 (2023). doi: 10.3969/j.issn.1672 0636.2023.04.007

[7] Y. Lyu, W. Wang and P. Chen, Fixed-Wing UAV based air-to-ground channel measurement and modeling at 2.7GHz in rural environment. IEEE T. Antenn. Propag. 73, 2038–2052 (2025). doi: 10.1109/TAP.2024.3428337

[8] J. Wang, Q. Zhu, Z. Lin et al., Sparse bayesian learning-based hierarchical construction for 3D radio environment maps incorporating channel shadowing. IEEE T. Wirel. Commun. 23, 14560–14574 (2024). doi: 10.1109/TWC.2024.3416447

[9] Y. Xiao, Q. Yuan, K. Jiang et al., Frequency-Assisted mamba for Remote Sens-Basel image super-resolution. 2024. doi: 10.48550/arXiv.2405.04964

[10] 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). doi:10.1016/j.knosys.2021.107678

[11] Y. Xu, X. Li, Y. Jie et al., Simultaneous Tri-Modal medical image fusion and super-resolution using conditional diffusion model. 2024. doi: 10.48550/arXiv.2404.17357

[12] C. Dong, C.C. Loy, K. He et al., Learning a deep convolutional network for image super-resolution. Computer Vision –ECCV 2014. ECCV 2014. Lecture Notes in Computer Science 8692. Springer, Cham. doi:10.1007/978-3-319-10593-2_13

[13] W. Shi, J. Caballero, F. Huszár et al., Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016. doi:10.1109/CVPR.2016.207

[14] J. Kim, J.K. Lee, K.M. Lee, Accurate image super-resolution using very deep convolutional networks. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016. doi:10.1109/CVPR.2016.182

[15] C. Ledig, L. Theis, F. Huszár et al., Photo-Realistic single image super-resolution using a generative adversarial network. IEEE Computer Society, 2016. doi:10.1109/CVPR.2017.19

[16] J.Y. Li, X.Y. Li, J. Zhang, based resolution reconstruction super-Image on SRGAN. Computer Knowledge Technology (2024). doi:10.14004/j.cnki.ckt.2024.0029

[17] Z.W. Zhou, Research on the super-resolution of engineering cost bills based on SRGAN. Intelligent City 9(10), 105–107 (2023). doi:10.19301/j.cnki.zncs.2023.10.033

[18] K. Zhang, W. Zuo and L. Zhang, Learning a single convolutional super-resolution network for multiple degradations. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2018. doi:10.1109/CVPR.2018.00344

[19] X.C. Hu, Research on adaptive image super-resolution algorithms. University of Sci. Technology of China, 2021. doi:10.27517/d.cnki.gzkju.2021.000009

[20] G.L. Wu, Research on blind image super-resolution algorithms based on deep learning. Xiamen University, 2021. doi:10.27424/d.cnki.gxmdu.2021.001765

[21] K. Zhang, J. Liang, L.V. Gool et al., Designing a practical degradation model for deep blind image super-resolution. 2021. doi:10.48550/arXiv.2103.14006

[22] Z. Luo, H. Huang, L. Yu et al., Deep constrained least squares blind image super-resolution. 2022. doi:10.48550/arXiv.2202.07508

[23] Y. Yuan, S. Liu, J. Zhang et al., Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks. IEEE, 2018. doi:10.1109/CVPRW.2018.00113

[24] 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). doi:10.1029/2020GL089029

[25] A. Liu, Y. Liu, J. Gu et al., Blind image super-resolution: A survey and beyond. IEEE T. Pattern Anal. 45, 5461–5480 (2022). doi:10.1109/TPAMI.2022.3203009

[26] A. Bulat, J. Yang and G. Tzimiropoulos, To learn image super-resolution, use a GAN to learn how to do image degradation first. European Conference on Computer Vision. Springer, Cham, 2018. doi:10.1007/978-3-030-01231-112

[27] M. Fritsche, S. Gu and R. Timofte, Frequency separation for real-world super-resolution. IEEE, 2019. doi:10.1109/ICCVW.2019.00445

[28] Y. Zhou, W. Deng, T. Tong et al., Guided frequency separation super-resolution. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, 2020. doi: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. doi:10.48550/arXiv.2107.10833

[30] X. Wang, K. Yu, S. Wu et al., ESRGAN: enhanced super-resolution generative adversarial networks. 2018. doi: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). doi:10.1016/0165-1684(89)90048-0

[32] M. Kirchner and J. Fridrich, On detection of median filtering in digital images. International Society for Optics and Photonics 7541, 754110 (2010). doi:10.1117/12.839100

[33] A. Ruangma, B. Bai, J.S. Lewis, Three-dimensional maximum a posteriori (MAP) imaging with radiopharmaceuticals labeled with three Cu radionuclides. Nuclear Medicine & Biology 33(2), (2006). doi: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). doi:10.1016/j.neuroimage.2005.10.037

[35] Y. Kataoka, T. Matsubara and K. Uehara, Image generation using generative adversarial networks and attention mechanism. IEEE/ACIS International Conference on Computer & Information Science. 2016. doi: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). doi:10.1007/s11760-020-01790-5

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv −Machine translation. Verify with original.