Image Super-Resolution via CNN-Guided Prior Integration in PDE-Based Reconstruction Framework

Tianfu Ji
Teaching Affairs Office, Lianyungang Technical College, Lianyungang 222006, China
E-mail: 13645132878@163.com

Keywords: partial differential equations, image super-resolution reconstruction, CNN, attention module, prior information

Received: May 20, 2025

The traditional image super-resolution reconstruction method has the problems of relying on hardware devices, high computational cost, and poor interpretability and generalization ability. To improve the efficiency of image super-resolution reconstruction, the study proposes a partial differential equation image super-resolution reconstruction method that introduces a priori information. The study first extracts the a priori information of the image based on convolutional neural network, and then fuses the extracted a priori information with the partial differential equation model. This convolutional neural network is based on the ResNet-18 framework. It enhances the differential expression of feature channels and the precise capture of edge features. This is achieved by removing batch normalization layers and introducing channel attention modules and gradient-guided branches. The experiment was conducted on the Flickr2K dataset and evaluated using cross method comparison metrics such as structural similarity index and peak signal to noise ratio. The results indicated that feature similarity and edge preservation rate were highest when extracting the gradient information of the image based on convolutional neural network when compared with other methods. When the number of iterations was 500, the feature similarity and edge preservation rate were 0.88 and 88.7% respectively. Edge pixel accuracy and gradient feature correlation were best when extracting gradient information from the image based on convolutional neural network. The values of edge pixel accuracy and gradient feature correlation were 0.92 and 0.87 respectively when the iteration was 500. The proposed method of partial differential equation image super-resolution reconstruction by introducing a priori information has superior performance and can provide technical support for image super-resolution reconstruction.

Povzetek: Razvita je metoda za izboljšanje ločljivosti slik, ki združuje konvolucijske nevronske mreže (CNN) z modelom delnih diferencialnih enačb (PDE). CNN, zasnovan na izboljšanem ResNet-18 brez normalizacijskih plasti, vključuje kanalno pozornost in gradientno vodenje za natančnejše zaznavanje robov. Iz mreže izluščene gradientne informacije služijo kot predhodno znanje, ki se vključi v PDE-model in usmerja proces rekonstrukcije.

1 Introduction

With the popularization of smartphones and the rise of social media, visual culture is becoming more and more dominant in modern society [1]. More and more people choose to obtain and share information through pictures. Social media platforms are dominated by images, and users share and consume images on these platforms far more frequently than text. This change has not only changed the way human beings acquire information, but also influenced their way of thinking, making visual expression a part of daily communication [2]. However, during image acquisition, the image resolution is often insufficient due to factors such as sensor shape and size, air disturbance, object motion, and lens defocusing, leading to loss of details and degradation of clarity, which in turn affects subsequent analysis and applications [3-4]. By using hardware or software to reconstruct the appropriate high resolution (HR) images from low resolution (LR) photos, the image super-resolution reconstruction (ISRR) technology has emerged as a successful solution to this issue. The aim of this technique is to increase the resolution of the image through algorithms without increasing the cost of hardware, thus obtaining an image that contains more information [5-6]. However, traditional ISRR methods rely heavily on hardware devices, involve complex computational processes, and struggle to meet real-time requirements. In addition, insufficient use of prior information in the image results in inadequate detail recovery and the introduction of artifacts. Therefore, the study proposes a partial differential equation (PDE) iterative step-response (ISR) method that introduces a priori information. This method constructs an innovative model for extracting a priori information based on convolutional neural networks (CNNs). It adopts an improved residual network (ResNet) structure by removing the batch normalization layer (BNL) to keep the color distribution consistent. The channel attention module (CAM) and gradient-guided branch are also introduced to enhance the differential representation of feature channels and the accurate capture

of edge features. This study aims to construct a prior information extraction model based on an improved CNN. The model introduces a CAM and a gradient guidance branch to enhance differential expression of the feature channels and accurately capture edge features. This enables effective reconstruction of LR images. The success criteria are based on 500 iterations with a structural similarity index (SSIM) greater than 0.95 and a peak signal-to-noise ratio (PSNR) greater than 40 dB to demonstrate the superior performance of the proposed method in ISRR.

2 Related works

CNNs are sophisticated graph-based representation models that have extensive application in a variety of domains [7-8]. In an attempt to reduce the pressure on experts and equipment work due to the increase in the number of patients with diabetic retinopathy, Alshawabkeh et al. The study proposed a hybrid CNN model that combines image enhancement, contrast limited adaptive histogram equalization, migration learning, and integrated classification techniques. The results indicated that the accuracy, precision, recall, and stability of the method proposed in the study were higher [9]. Xiong et al. proposed a molecular CNN architecture based on DNA regulatory circuits to address the limitations of traditional neural networks in biomolecular recognition. The study combined DNA molecular circuits with deep learning (DL) to construct a novel neural network model with molecular computational properties. The findings demonstrated that, in comparison to the conventional approach, the suggested method's recognition accuracy was greatly increased to 98.7% [10]. A hybrid CNN-based model was proposed by Gupta et al. to handle the difficulties of picture quality, dataset imbalance, and dataset generation from various sources. The study combined three separate base hybrid CNN models in parallel configurations to offset the drawbacks of individual models. With an overall test accuracy of 97.3%, the study's suggested hybrid model beaten the majority of models, according to the data [11]. Çelik et al. proposed a hybrid CNN model and created a new deep feature in order to increase the dataset of durum wheat seeds for recognition and classification. The study classified the new feature set as support vector machine input. The results showed that the study proposed a model to recognize and classify durum wheat and a new durum wheat dataset was obtained [12]. A hybrid CNN model based on maximum correlation minimum redundancy was presented by Eroglu et al. to address the issue of Alzheimer's disease not being identified and categorized early enough to successfully delay the disease. The study classified signs in magnetic resonance imaging of the brain. The results displayed that that the accuracy of the risen model for feature extraction (FE) and classification was improved to 99.1% [13].

The imaging environment, imaging distance, optical system error, and other factors will all have an impact on the quality of the image during the acquisition process. When these elements are combined, the image quality will deteriorate. In contrast, super-resolution reconstruction (SRR) can reconstruct the corresponding HR image from the observed LR image [14]. To solve the issue of medical image resolution issues that impact clinical diagnosis accuracy, Du W et al. suggested an SRR technique that combines Transformer and generative adversarial networks. The study realized high-quality reconstruction of medical images by fusing the global FE capability of Transformer and the detail generation advantage of generative adversarial networks. The findings revealed that the suggested technique preserved the image's anatomical validity while increasing reconstruction accuracy to 98.3% [15]. Afacan et al. suggested a scan-specific generative neural network-based technique to enhance magnetic resonance imaging resolution and produce high-quality image reconstruction. The DL algorithm was used to perform SRR of LR magnetic resonance images. The outcomes demonstrated how well the suggested technique improved image detail reproduction. In terms of PSNR and SSIM, the reconstructed image performed noticeably better than the traditional approach [16]. To enhance ISRR performance, Zhang M et al. proved a DL network based on heat transfer theory. The results displayed that that this method achieved better reconstruction results than the traditional algorithm on several benchmark datasets [17]. Zhang et al. proposed a SRR method for the problem of insufficient resolution of global 3-arc-second digital elevation model data. The study constructed a DL-driven digital elevation model SRR framework by fusing multi-source remote sensing data. The results indicated that the proposed method successfully improved the resolution of the original digital elevation model data to the level of 1 arc second, and the error of elevation accuracy was controlled within ±2.5 meters [18]. To meet the needs of modern video processing, Gong et al. proposed a video SSR method based on the Transformer and attention mechanisms. They also designed a video super-resolution architecture based on the temporal Transformer. The results indicated that the proposed model had substantially improved image quality [19]. The related works summary table is shown in Table 1.

Table 1: Summary table of related works.

Literature	Method	Data set	Index	Limitation
[9]	Hybrid CNN+Transfer Learning+Ensemble Classification	Retinal fundus image	Accuracy, precision, and recall have all been improved	Dependent on image quality, without specifying generalization ability
[10]	CNN	Synthetic DNA Sequence Dataset	The recognition accuracy is 98.7%	Scalability is limited
[11]	Parallel Hybrid CNN	COVID-19 public dataset	The accuracy rate is 97.3%	There is an issue of data imbalance

[12]	Hybrid CNN+SVM classification	Self built hard grain wheat grain dataset	-	Small dataset size
[13]	Hybrid CNN	ADNI Public Dataset	The classification accuracy is 99.1%	Relying on MRI quality
[15]	Transformer+GAN	Medical Imaging Dataset	The reconstruction accuracy is 98.3%	High consumption of computing resources
[16]	Scan specific generative neural network	Low resolution MRI	SSIM and PSNR have both been improved	Generalization limited by scanning protocol
[17]	Heuristic deep learning network	Set5, Set14	PSNR increased by 1.2dB, SSIM increased by 0.03	Simplification of physical models leads to loss of high-frequency details
[18]	Deep Learning Framework	Global 3-second DEM data	Resolution increased to 1 arc second	Increased errors in complex terrain areas
[19]	Transformer+cross modal attention mechanism	Video dataset	Substantial improvement in image quality	Time consistency indicator not specified

To summarize, ISRR technique is of great significance for improving image quality in medical diagnosis, remote sensing mapping and other fields. To further increase ISRR's accuracy and dependability, numerous researchers and experts have created numerous enhanced models. However, there are still some shortcomings, such as limited adaptability to complex imaging environments and low computational efficiency in processing special images. Therefore, the study proposes a PDE ISRR method that incorporates a priori information. This method aims to improve the accuracy and stability of SRR and reduce the blurring caused by traditional methods.

3 Introduction of a priori information for PDE ISRR

3.1 DL-based a priori information extraction models

In DL, CNN is a strong network structure, particularly for processing images and videos [20]. CNNs' primary strength lies in their ability to automatically and effectively extract features from data, a task that is often done manually in conventional machine learning techniques. CNNs are designed with convolutional layers (CLs) so that each neuron only needs to respond to a portion of the input data, i.e., the local perceptual domain. This mechanism allows CNNs to capture local features in an image, such as edges, textures, etc [21-22]. Therefore, the study employs CNN for FE of LR images to learn the gradient, texture and other information of the image, which is input into the PDE model as a priori information. The retrieved features are more abstract and include more semantic information the more layers the CNN network has. Nevertheless, when the layer of the model increases, it leads to the problem of gradient dispersion or gradient explosion. To solve this problem, the research adds ResNet to the a priori information extraction model. In ResNet, the output of each sublayer is not just the result of the output of the previous layer after an activation function (AF). However, it is directly added to the input of the previous layer through jump connections. Suppose there is a layer with input x and output $F(x)$ after a series of transformations, the final output is displayed in Equation (1).

$$y = F(x) + x \tag{1}$$

In Equation (1), y is the result of residual linkage. With this design, the model no longer needs to learn the entire input representation, but rather the increment of the input. In the SRR task, it is crucial to maintain the consistency between the input and output images in terms of color distribution. In contrast, the BNL in the ResNet architecture changes the distributional properties of the input data through normalization operations. This process may interfere with this consistency, leading to color distortion or contrast anomalies in the reconstructed images. Therefore, in the design of the a priori information extraction model, the removal of the BNL in ResNet is investigated. The ResNet pairs before and after removal are shown in Figure 1.

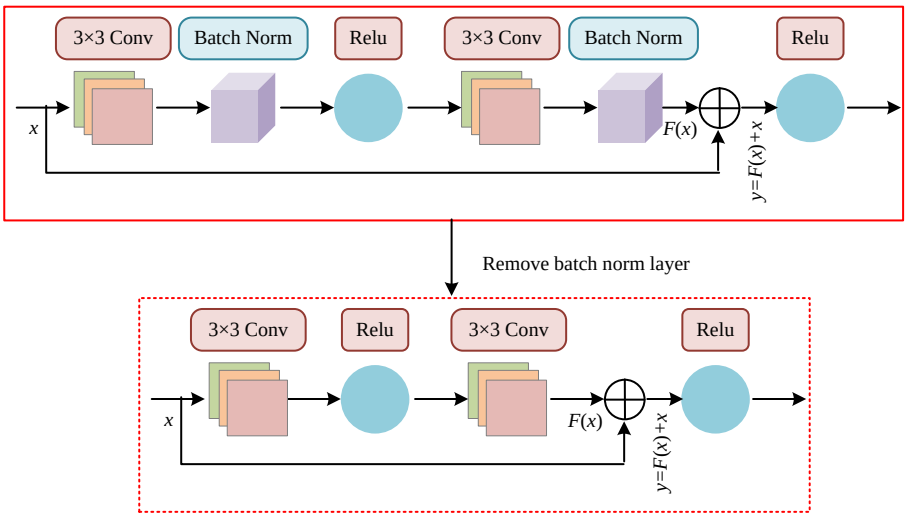


Figure 1: Comparison of ResNet before and after removing BNL.

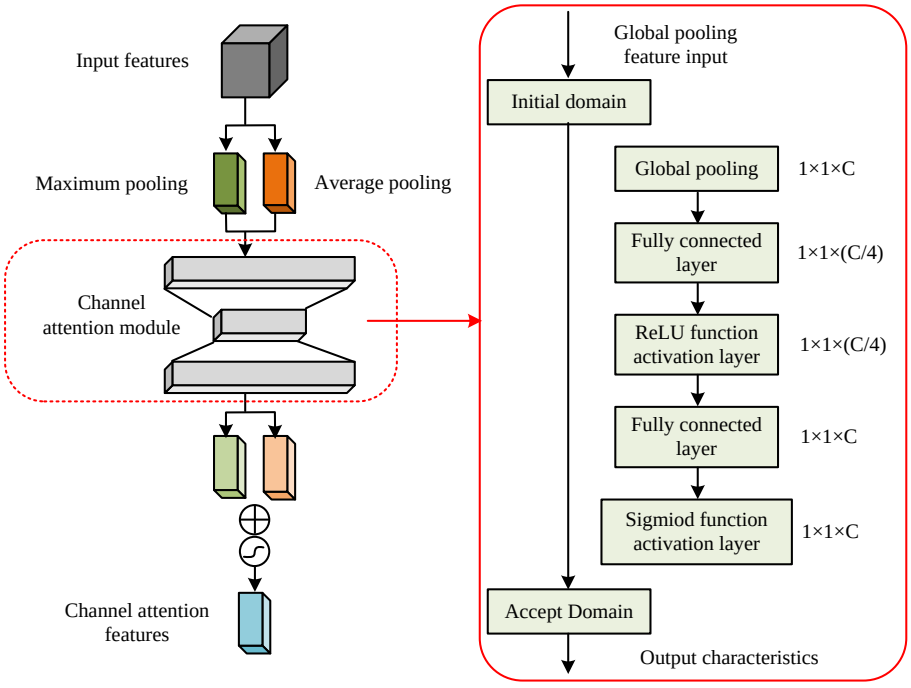


Figure 2: Channel attention calculation steps.

In Figure 1, the upper part shows the original ResNet structure containing two consecutive 3×3 CLs. The BNL and ReLU AF are connected after each CL. The lower part shows the improved structure after removing the BNL, which retains the 3×3 CLs and ReLU AF and removes all the BNLs. To improve the network's ability to perceive key regions of the image, especially the reconstruction effect in detail-rich regions and edge structures. The study incorporates two core modules in the model design. Among them, the CAM is used to enhance the differential representation between feature channels. The gradient guidance branch is used to capture and reconstruct the edge features of the image. The channel attention (CA) computation steps are shown in Figure 2.

In Figure 2, the steps of CA computation are as follows. First, global maximum pooling (GMP) and global average pooling (GAP) of spatial dimensions are performed on an input feature map (FM) F of size H×W×C to obtain two 1×1×C FMs. The GAP is computed in Equation (2).

$$F_{avg} = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} X_{ij} \tag{2}$$

In Equation (2), F_{avg} is the FM after GAP. X_{ij} is the value of the input FM at position (i, j). W and H is the width and height of the FM. The GMP is calculated in Equation (3).

$$F_{max} = \max_{i,j} X_{ij} \tag{3}$$

In Equation (3), F_{max} denotes the FM after GMP. Next, input F_{avg} and F_{max} into two shared multi layer perceptrons (MLPs) for learning, resulting in two feature

maps MLP_{avg} and $MLP_{\max}$ of 1×1×C. The calculation of MLP_{avg} is shown in equation (4).

$$MLP_{avg} = \text{ReLU}(W_1 F_{avg} + b_1) \tag{4}$$

In Equation (4), W_1 denotes the weight matrix (WM) of the first layer. b_1 denotes the bias vector of the first layer. ReLU denotes the ReLU AF. The calculation of $MLP_{\max}$ is shown in Equation (5).

$$\text{MLP}_{\max} = W_2 \text{MLP}_1 + b_2 \tag{5}$$

In Equation (5), W_2 is the WM of the 2ndL. b_2 is the bias vector of the 2ndL. Finally, the MLP output is subjected to addition operation and mapped by Sigmoid AF to obtain the final CA WM. The CA WM A is calculated in Equation (6).

$$A = \sigma(MLP_1 + MLP_2) \tag{6}$$

In Equation (6), σ denotes the Sigmoid AF. Gradient branching aims to super-resolve the gradient map (GM) of an LR image into the corresponding GM of an HR image. The GM of an image I is obtained by calculating the difference between neighboring pixels. Equation (7) provides the computation of the gradient vector.

$$\nabla I(\mathbf{x}) = (I_x(\mathbf{x}), I_y(\mathbf{x})) \tag{7}$$

In Equation (7), $\nabla I(\mathbf{x})$ is the gradient vector at position $\mathbf{x}$. $I_x(\mathbf{x})$ and $I_y(\mathbf{x})$ is the gradient along the x and y direction at position $\mathbf{x}$. The GM of image I is computed in Equation (8).

$$G(I) = \left\|\nabla I(\mathbf{x})\right\|_2 = \sqrt{I_x^2(\mathbf{x}) + I_y^2(\mathbf{x})} \tag{8}$$

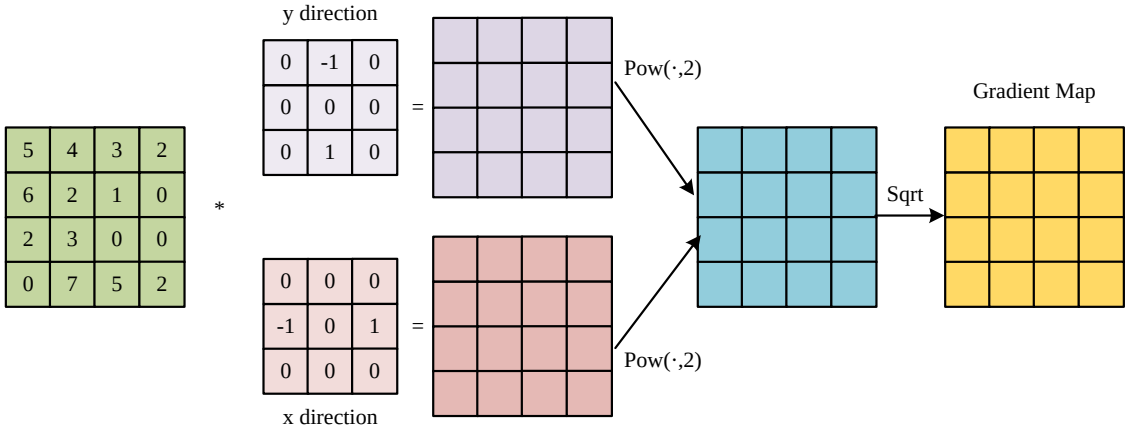


Figure 3: Steps for obtaining image gradients through CLs.

In Equation (8), $G(\cdot)$ denotes the operation of taking the GM, and the operation of obtaining the gradient can be realized by a CL with a fixed kernel. The realization steps are shown in Figure 3.

In Figure 3, the steps for obtaining the gradient of an image by means of a CL are as follows. First, two distinct convolution kernels are used to the image matrix in an attempt to determine the image's gradient in the x and y directions, respectively. After the convolution operation is completed, the gradient of the image in x and y direction is obtained. Next, each element of these two gradient matrices is squared and the results are summed to get a new matrix. Ultimately, the square root of this matrix yields the final GM. In summary, the CNN architecture proposed in this study is based on the ResNet-18 framework. It includes an initial CL with 64 3x3 convolution kernels and a ReLU AF. This is followed by 18 improved residual blocks. Each residual block contains two CLs. Each layer uses 64 3x3 convolution kernels and a ReLU AF. The input is added directly to the output of the second CL through residual connections. This study removes the BNL from the residual block and introduced a CAM. This module performs both global average and GMP on the input feature map simultaneously, producing two $1 \times 1 \times C$ vectors. Through shared two-layer MLP processing, they are added and activated by Sigmoid to generate channel weights. No pre trained model is used, and all network parameters are trained from scratch using the Adam optimizer.

3.2 PDE model construction and solution with fused a priori

The study first extracts the a priori information of the image based on CNN after which the extracted a priori information is fused with the PDE model. Most of the semantic and shape information in an image can be represented by edges. The edge portion of the image is where the pixel values change drastically. Gradient of all the pixel value locations of the image tells which locations in the image are edges. The gradient information (GI) essentially describes the trend of the pixel values, such as the contour of the object, the direction of the texture, and so on. These features are crucial for reconstructing HR images. Therefore, the study fuses the image GI extracted by the CNN model as a priori information with the PDE model. In ISRR, using GI as a priori information can help the reconstruction algorithm to recover the details of the image more accurately. This is particularly true in high-frequency areas, which typically hold the image's texture and edge information. The reconstructed HR image can be made to match the LR image at the pixel level by including this a priori information into the PDE model. Additionally, it can enhance the overall quality of the reconstructed image by lowering potential artifacts and blurring throughout the reconstruction process. The model that

integrates the image GI extracted by the CNN model as a priori information with the PDE is shown in Figure 4.

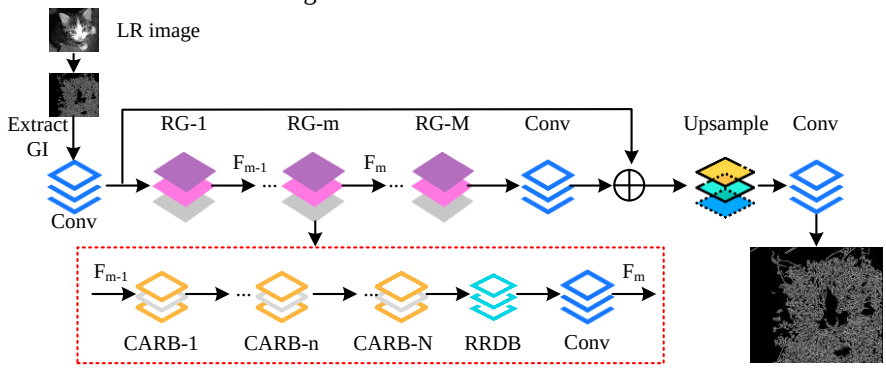


Figure 4: The image GI extracted from the CNN model as a model for integrating a priori information with PDEs.

In Figure 4, the model is mainly composed of an input layer, a CL, a ResNet layer, an MLP, a feature fusion layer, and an output layer. First, the initial LR map is passed through a CL to extract the preliminary FM. Second, the preliminary FMs go to the ResNet layer, where the features are further extracted and optimized. Then, the extracted feature maps are fused through a feature fusion layer to form a fused feature map. Subsequently, the fused feature map is input into the PDE model and fused with the GI image extracted from the CNN model. The PDE model utilizes GI as prior information to guide the SSR process of images. Next, the fused feature map is then input into the PDE model, where it is fused with the GI image extracted from the CNN model. Finally, the upsampled image is refined through a CL to obtain the final high-resolution image.

Finite difference method (FDM) is a numerical method for solving PDEs and ordinary differential equations. Its solved on computer by discretizing differential equations into difference equations. It has the advantages of simplicity and intuition, versatility and accuracy [23-24]. Therefore, this study is based on FDM for solving PDEs with fused a priori. FDM works on the basis of first breaking down the problem's definition domain in a grid. To simplify the PDE definite solution problem with continuous variables to a system of algebraic equations with just a finite number of unknowns, the derivative is substituted by the difference quotient of the function at the grid points [25-26]. The solution steps are shown in Figure 5.

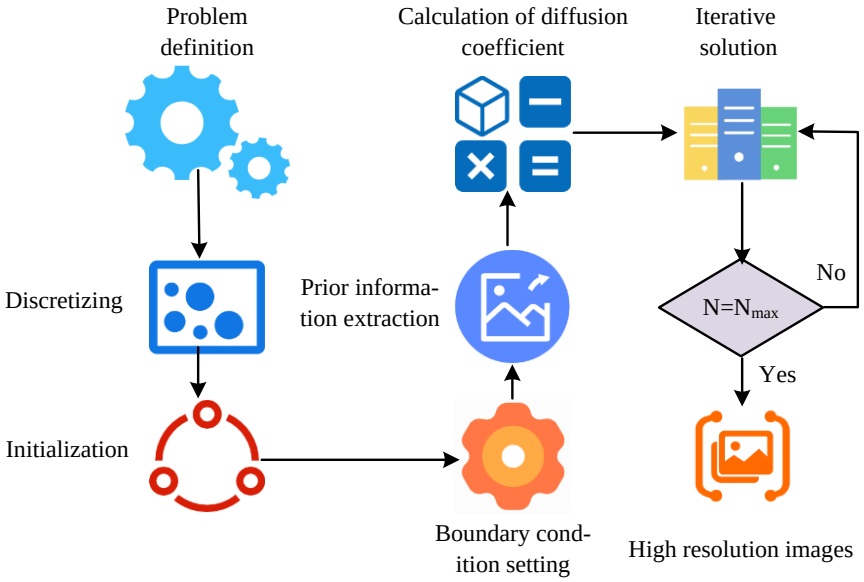


Figure 5: Solves the steps of the PDE of the fusion prior based on FDM.

In Figure 5, the steps for solving the PDE with fused prior are as follows. First, the problem domain is defined and the solution region is discretized in space and time to form a grid. Second, the initial and boundary conditions are set. The LR image's a priori information is then retrieved using CNN, and each grid point's diffusion coefficient is computed using the information that is extracted. Finally, the a priori information is fused and the features learned from the data are integrated into the PDE model to obtain the HR image. However, FDM is prone to numerical oscillations in high-resolution image reconstruction, especially when dealing with complex image boundaries and regions with high gradients. This can lead to image distortion. Moreover, high-resolution reconstruction requires finer grids to improve accuracy, which significantly increases the computational and

storage requirements of FDM. Larger grid sizes can also lead to insufficient accuracy. Therefore, this study adopts grid-adaptive encryption technology in complex areas and near image boundaries to refine the grid and capture details and boundary features more accurately. At the same time, the interpolation method is optimized to reduce the error introduced due to the mismatch between the grid and the image boundary. FDM substitutes a finite quantity of discrete points for the independent variable's continuous variation region. It substitutes functions of discrete variables defined on the grid points for functions of continuous variables that appear in the problem. The PDE discretization is shown in Equation (9).

$$\frac{u_{i,j}^{n+1}-u_{i,j}^{n}}{\Delta t}=D\left(\frac{u_{i+1,j}^{n}-2u_{i,j}^{n}+u_{i-1,j}^{n}}{\Delta x^{2}}+\frac{u_{i,j+1}^{n}-2u_{i,j}^{n}+u_{i,j-1}^{n}}{\Delta y^{2}}\right)+f_{i,j}^{n} \tag{9}$$

In Equation (9), $u_{i.j}^{n}$ denotes the pixel value at time step n and spatial location (i,j). D denotes the diffusion coefficient. Δt denotes the time step. Δx and Δy denote the spatial step. $f_{i.j}^{n}$ denotes the external source term. Among them, $\frac{u_{i,j}^{n+1}-u_{i,j}^{n}}{\Delta t}$ represents the rate of change over time. It reflects how image pixel values evolve with time steps. This enables the model to gradually optimize the image's details and structure based on prior information. $D\left(\frac{u_{i+1,j}^{n}-2u_{i,j}^{n}+u_{i-1,j}^{n}}{\Delta x^{2}}+\frac{u_{i,j+1}^{n}-2u_{i,j}^{n}+u_{i,j-1}^{n}}{\Delta y^{2}}\right)$ describes the spatial diffusion process of image pixel values. This process can smooth out noise and discontinuities in an image while enhancing edge and texture details. The result is a clearer, more complete image. $f_{i.j}^{n}$ may contain prior information extracted from CNN. This enables the model to utilize prior knowledge of the image better, restoring the details and structure of high-resolution images. This improves the quality and accuracy of reconstruction. The calculation of the diffusion coefficient is shown in Equation (10).

$$D=\alpha\cdot\exp(\beta\cdot|\nabla u|^{2}) \tag{10}$$

In Equation (10), α denotes the scaling factor. β denotes the adjustment coefficient. $|\nabla u|^{2}$ denotes the square of the mode of the image gradient. The gradients along the x and y directions are calculated in Equation (11).

$$\begin{cases}\frac{\partial u}{\partial x}=\frac{u_{i+1,j}^{n}-u_{i-1,j}^{n}}{2\Delta x}\\ \frac{\partial u}{\partial y}=\frac{u_{i,j+1}^{n}-u_{i,j-1}^{n}}{2\Delta y}\end{cases} \tag{11}$$

In Equation (11), $\frac{\partial u}{\partial x}$ is the gradient along the x direction. $\frac{\partial u}{\partial y}$ is the gradient along the y direction. The boundary conditions are shown in Equation (12).

$$u_{0,j}^{n}=u_{N_x,j}^{n}=u_{i,0}^{n}=u_{i,N_y}^{n}=g_{i,j} \tag{12}$$

In Equation (12), $g_{i,j}$ denotes the boundary condition value. In ISRR, it is often difficult to obtain the values of boundary pixels directly from LR images. Equation (12) specifies the values of the image's boundary pixels, providing a constraint condition for the image's edges. This ensures the image's rationality at the boundary and the accuracy of internal pixel calculations. The iterative update is shown in Equation (13).

$$u_{i,j}^{n+1}=u_{i,j}^{n}+\Delta t\cdot\left(D\left(\frac{u_{i+1,j}^{n}-2u_{i,j}^{n}+u_{i-1,j}^{n}}{\Delta x^{2}}+\frac{u_{i,j+1}^{n}-2u_{i,j}^{n}+u_{i,j-1}^{n}}{\Delta y^{2}}\right)+f_{i,j}^{n}\right) \tag{13}$$

PSNR measures the difference between a reconstructed image and the original image, reflecting the similarity of the pixel values in the image. Higher PSNR values indicate better image quality. The calculation is shown in equation (14).

$$PSNR=10\cdot\log_{10}\left(\frac{MAX^{2}}{MSE}\right) \tag{14}$$

In equation (14), MAX represents the maximum possible value of an image pixel, and MSE represents the maximum possible value of an image pixel. The SSIM is used to evaluate the similarity between reconstructed and original images. It takes into account brightness, contrast, and structural information. The range of SSIM values is between -1 and 1, and the closer the value is to 1, the more similar the image structure is. This study uses functions in MATLAB to directly calculate PSNR and SSIM metrics.

4 Quality analysis of SRR of PDE images

4.1 Experimental environment and parameter settings

The hardware configuration chosen for the experiment is as follows. The operating system is Windows 11, the RAM is 64GB, the video memory is 24GB, the CPU is Intel Core i9-13900K @3.00GHz, and the GPU chosen is NVIDIA-GeForce RTX 4090. The software chosen for this study is yTorch 1.12, CUDA 11.6, and MATLAB-R2023a. This study conduct experiments by using the Flickr2K dataset (http://cv.snu.ac.kr/research/EDSR/Flickr2K.tar). The Flickr2K dataset is an image dataset used for super-resolution tasks, containing 2650 high-resolution images. First, the image is normalized by reducing the pixel values

to within the range of [0,1]. In addition, in order to enhance the generalization ability of the model, data augmentation processing is performed on the data, including random cropping, horizontal flipping, vertical flipping, and rotation, to increase the diversity and richness of the training data. It is divided into training set, test set, and validation set in the ratio of 8:1:1. Table 2 displays the experimental parameters' precise settings.

Table 2: Sample parameter settings.

The parameter name	Parameter values	Describe	The parameter name	Parameter values	Describe
Enter image size	64×64	Input dimensions for the low-resolution images	BatchSize	16	Number of images per training input
Learning rate	1×10-4	Initial learning rate of the Adam optimizer	Epochs	500	Number of training iterations for the complete dataset
ResNet number of plies	18	Number of underlying layers of the residual network	Gradient branch convolution kernel	3×3 Center Difference	Convolution kernel fixed for extracting the image ladder, degrees
Channel attention dimension	64	The characteristic dimension of the MLP middle layer	PDE coefficient of diffusion	0.05	Controlling for the weight of the gradient prior in the PDE model
(△x,△y)	0.5	Spatially discretized the grid step size	△t	0.01	The time step of the PDE iterative solution

Table 3: Parameter sensitivity analysis results.

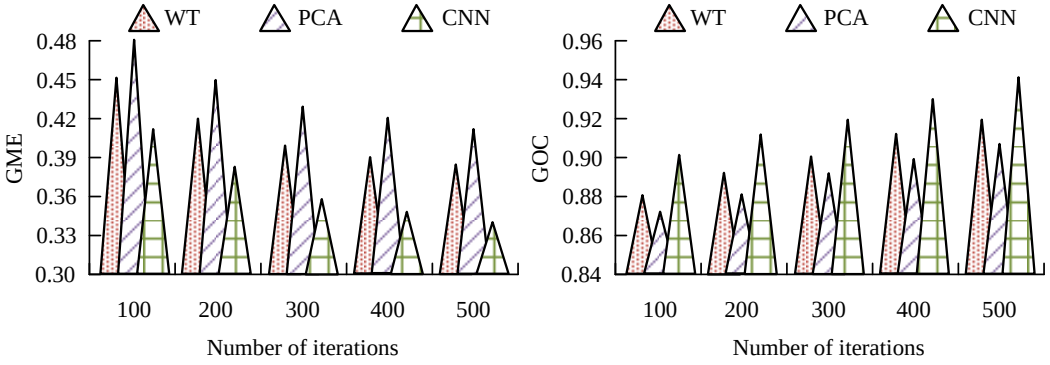
α	β	PSNR/dB	SSIM	AI
0.03	0.1	40.45±0.29	0.95±0.01	0.55±0.07
0.05	0.1	40.87±0.22	0.96±0.01	0.51±0.04
0.08	0.1	41.43±0.31	0.96±0.01	0.49±0.05
0.03	0.2	40.78±0.21	0.96±0.01	0.53±0.05
0.05	0.2	41.75±0.17	0.97±0.01	0.47±0.04
0.08	0.2	40.38±0.27	0.95±0.01	0.59±0.08
0.03	0.3	40.21±0.25	0.93±0.01	0.52±0.05
0.05	0.3	41.19±0.23	0.96±0.01	0.59±0.07
0.08	0.3	40.87±0.25	0.95±0.01	0.60±0.09

α is used to control the weight of gradient priors in PDE models. A larger α can highlight image edges and textures, but may also amplify noise. A smaller α may lead to blurred edges. β is used to regulate the sensitivity of diffusion coefficient to gradient changes. A larger β makes the diffusion coefficient more sensitive to gradient changes. This highlights strong edges but potentially loses weak textures. A smaller β is the opposite. To determine the parameters α and β in the diffusion coefficient, the study set α to 0.03, 0.05, and 0.08, and set β to 0.1, 0.2, and 0.3, respectively. PSNR, SSIM, and artifact index (AI) are used as evaluation metrics. The results of parameter sensitivity analysis are shown in Table 3. When α =0.05 and β =0.2, all indicators reach their optimal values. This indicates that the parameter combination can achieve good results in ISRR.

4.2 Quality analysis of CNN-based a priori information extraction models

The qualitative judgment of reconstruction quality is based on domain consensus: PSNR not less than 40dB is excellent. SSIM not less than 0.95 is high fidelity. AI not more than 0.5 is no obvious artifacts. Local texture sharpness (LTS) not less than 0.8 is clear texture. SDR not more than 0.1 is structural fidelity. Gradient magnitude error (GME) and gradient orientation consistency (GOC) are compared with other methods for extracting the GI of an image based on CNN with different number of iterations. GME is evaluated by calculating the absolute error of the gradient amplitude between the reconstructed image and the original image. The smaller the value, the better the consistency of the gradient amplitude. The GOC evaluates the consistency of directions by calculating the difference in the gradient directions between reconstructed and original images. It can also use the cosine similarity of the angle differences to calculate the consistency of directions. The closer the GOC value is to 1, the better the consistency of the gradient direction. The comparison of GME and GOC for extracting image GI using different methods is shown in Figure 6. In Figure 6(a), the GME decreases with the increase in the number of iterations when different methods are used to extract the GI of the image. When CNN is used to extract the image's GI, the GME is at its lowest, and when principal component analysis (PCA) is employed, it is at its highest.

When the iteration is 500, the GME is categorized as 0.34 and 0.41. In Figure 6(b), the GOC increases with the increase in the number of iterations for extracting the GI of the image by different methods. The GOC is maximum when CNN is used to extract the GI of the image. When the number of iterations is 500, the GOC at this time is 0.94. The results display that CNN can extract the GI more accurately while maintaining higher directional consistency.

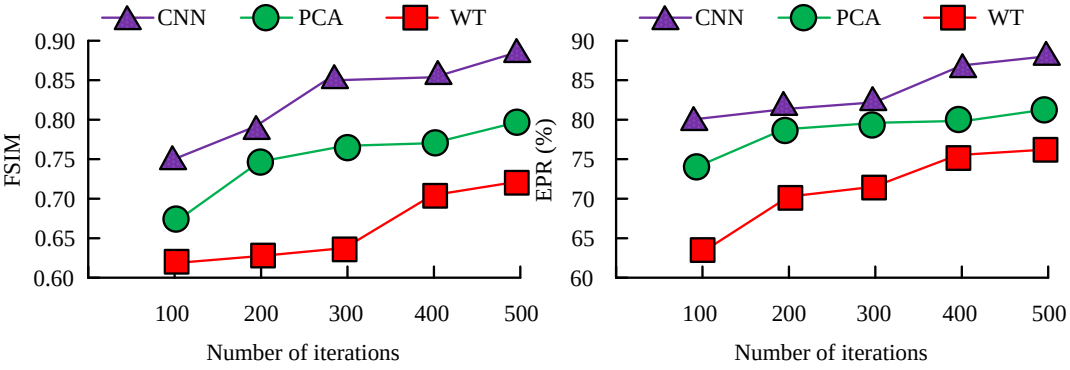


(a) GME for extracting gradient information from images using different methods

(b) GOC for extracting gradient information from images using different methods

Figure 6: Comparison of GME and GOC for extracting image GI using different methods.

With different number of iterations, the feature similarity (FSIM) and edge preservation rate (EPR) when extracting GI from the image based on CNN is compared with other methods. FSIM evaluates FSIM by calculating the gradient magnitude and gradient orientation of an image. The closer the value is to 1, the higher the FSIM of the image. The EPR calculates the retention rate by comparing the edge maps of the original and reconstructed images. These images can be obtained using the Canny operator or other edge detection methods. The comparison of FSIM and EPR for extracting image GI using different methods is shown in Figure 7. The higher the EPR value, the more edge information is preserved in the reconstructed image. In Figure 7(a), FSIM is lowest when using wavelet transform (WT) and greatest when using CNN to extract an image's GI. When the number of iterations is 500, the FSIM is 0.71 and 0.88, respectively. In Figure 7(b), the EPR is highest when the GI of the image is extracted based on CNN. When the number of experiments is 500, the EPR is 88.7%. It verifies the efficiency of the proposed method of the study in edge information retention.



(a) FSIM for extracting gradient information from images using different methods

(b) EPR for extracting gradient information from images using different methods

Figure 7: Comparison of FSIM and EPR for extracting image GI using different methods.

The edge pixel accuracy (EPA) and gradient feature correlation (GFC) when extracting GI from an image based on CNN is compared with other methods. The EPA calculates accuracy by comparing the edge maps of the reconstructed and original images. This calculation is performed by dividing the number of correctly classified edge pixels by the total number of edge pixels. The higher the EPA value, the higher the matching degree between the edge pixels of the reconstructed image and the original image. The GFC measures the correlation between the direction and magnitude of image gradients. The Pearson correlation coefficient can then be used to calculate the correlation of the gradient features. The higher the GFC value, the better the consistency of gradient features. Table 4 displays the findings. CNN based extracting the GI of an image shows good performance. Poor performance is shown when extracting GI of an image based on WT. The values of EPA are 0.92 and 0.87 when the iteration is 500. The values of GFC are 0.87 and 0.79. The reliability of CNN in edge localization and GFC is confirmed.

Table 4: EPA and GFC for ISRR using different models.

Number of iterations	EPA			GFC		
	WT	PCA	CNN	WT	PCA	CNN
100	0.79	0.83	0.85	0.72	0.75	0.79
200	0.80	0.85	0.87	0.74	0.77	0.81
300	0.82	0.86	0.89	0.75	0.79	0.84
400	0.84	0.88	0.91	0.77	0.80	0.85
500	0.87	0.89	0.92	0.79	0.83	0.87

Table 5: Results of ablation experiment.

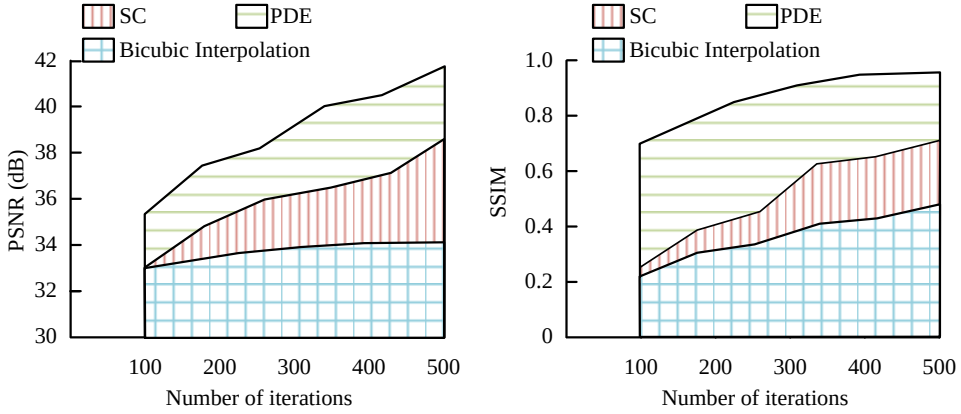
Model	PSNR/dB	SSIM	AI
CNN	38.42	0.91	0.62
CNN+CAM	39.25	0.93	0.58
PDE	37.89	0.89	0.66
Complete model	41.74	0.97	0.48

4.3 Fusion of a priori PDE models for quality analysis

Ablation studies are conducted to evaluate the effectiveness of each component in the proposed model. PSNR, SSIM, and AI are used as evaluation metrics to compare CNN, CNN+CAM, PDE, and the complete model. The results of the ablation experiment are shown in Table 5. The complete model performs the best in terms of metrics, with PSNR, SSIM, and AI being 41.74dB, 0.97, and 0.48, respectively. Next is CNN+CAM. PDE performs the worst, indicating that relying solely on PDE for reconstruction can easily result in the loss of detailed image information. The results suggest that combining the CNN, CAM, and PDE modules can enable the model to reconstruct images with different textures effectively.

To verify the effectiveness of fusing a priori PDE models for ISRR, the study uses PSNR with SSIM for evaluation. In Figure 8(a), the PSNR value of ISRR based on fusion a priori PDE model is the highest and the PSNR value based on bicubic interpolation is the lowest. When the number of iterations is 500, the PSNR values of bicubic interpolation and PDE model are 33.8dB and 41.74dB, respectively. In Figure 8(b), the highest SSIM value is obtained for SRR of the image based on the fused a priori PDE model. The SSIM value is 0.97 when the number of iterations is 500. The results displays that the proposed method of the study can effectively restore the image structure.

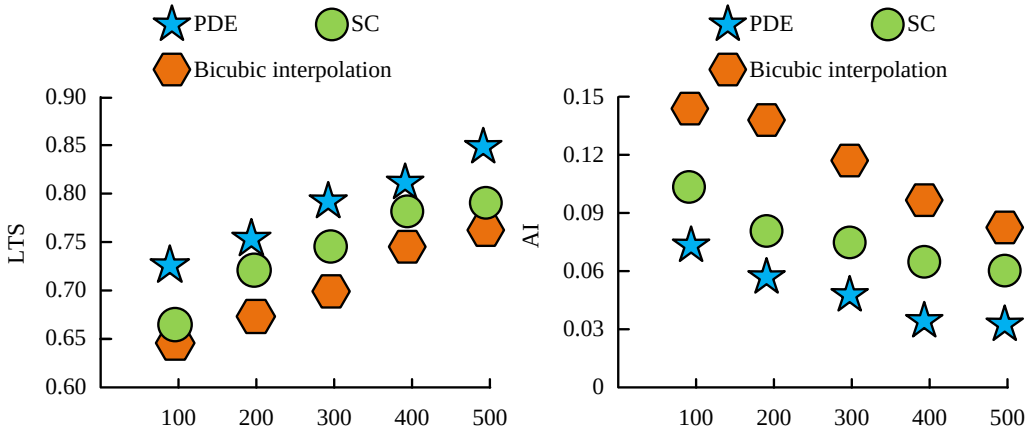
The LTS and AI of the SRR of the image by fusing the a priori PDE model are compared with other methods. LTS measures by calculating the contrast and sharpness of local regions in an image, with higher values indicating clearer texture details in the image. AI can evaluate the extent of artifacts in an image by calculating the distribution and intensity of abnormal pixels, with lower values indicating fewer artifacts. The comparison between LTS and AI for ISRR using different models is shown in Figure 9. In Figure 9(a), the LTS of ISRR by different methods increases with the iterations' quantity. The highest value of LTS is obtained for SRR of image based on fused a priori PDE model. When the iteration is 500, the LTS value is 0.852. In Figure 9(b), the AI of SRR of the image by different methods decreases with the increase in the iteration. The AI value of SRR of image based on fused a priori PDE model is minimum. When the iteration is 500, the AI value is 0.48. It has been confirmed that the study's suggested approach greatly enhances texture details while lowering artifacts.



(a) PSNR of image super-resolution reconstruction using different methods

(b) SSIM of image super-resolution reconstruction using different methods

Figure 8: PSNR and SSIM for ISRR using different models.



(a) LTS of image super-resolution reconstruction using different methods

(b) AI of image super-resolution reconstruction using different methods

Figure 9: Comparison between LTS and AI for ISRR using different models.

The edge enhancement index (EEI) and structural distortion rate (SDR) of ISRR by fusing the a priori PDE model are compared with other methods. EEI evaluates the edge enhancement effect by comparing the edge intensity of the reconstructed image with the original high-resolution image. A higher EEI value indicates a greater enhancement of the reconstructed image's edge intensity compared to the original image and a better edge detail restoration effect. SDR evaluates the degree of distortion in structural features between reconstructed and original images by comparing their local structural similarities and differences. The SDR value ranges from 0 to 1. A smaller value indicates a higher SSIM between the reconstructed and original images and a lower degree of distortion. Table 6 displays the findings. The fusion a priori PDE model shows good performance in ISRR. The bicubic interpolation based ISRR shows poor performance. With 500 iterations, the EEI values for the PDE model and bicubic interpolation are 0.64 and 0.54, respectively, while the SDR values are 0.05 and 0.10, respectively. The efficiency of the PDE model for edge enhancement and structure fidelity is verified.

To further validate the generalization and efficiency of the proposed model, the DIV2K dataset containing 1000 different scene images is used for testing. Evaluation metrics include PSNR, runtime, GPU utilization, and model size. The PDE model integrating a priori information is compared with three advanced SSR models: bicubic interpolation, super solution generative adversarial network (SRGAN), and hybrid attention Transformer (HAT). Thirty independent runs are conducted to collect data, and paired t-tests are used to evaluate whether the performance differences between the models are statistically significant. The threshold for significance level is 0.05. If $p<0.05$, it is considered that the performance improvement of the proposed method is significant. The performance comparison results of the four models are shown in Table 7. The average PSNR of the PDE model that incorporates prior information is 33.46 dB, which is significantly higher than that of the comparison model ($p<0.05$). With an average GPU utilization of 70.52% and an average running time of 0.66 s, the PDE model surpasses all but the bicubic interpolation model. The results indicate that the PDE

model incorporating prior information has excellent generalization and efficiency in ISRR tasks.

Table 6: EEI and SDR for ISRR using different models.

Number of iterations	EEI			SDR		
	Bicubic Interpolation	SC	PDE	Bicubic Interpolation	SC	PDE
100	0.42	0.50	0.55	0.17	0.15	0.11
200	0.47	0.51	0.58	0.15	0.14	0.09
300	0.50	0.53	0.59	0.13	0.12	0.07
400	0.52	0.56	0.61	0.11	0.10	0.06
500	0.54	0.58	0.64	0.10	0.09	0.05

Table 7: Performance comparison results of four models.

Model	PSNR/dB	Run time/s	Utilization rate/%	Model size/MB	p (and PSNR of PDE)
Bicubic Interpolation	28.55±0.33	0.15±0.03	15.20±2.07	0.55	<0.05
SRGAN	30.24±0.46	0.88±0.13	65.14±5.01	119.82	<0.05
HAT	31.88±0.39	0.66±0.09	69.82±6.25	85.43	<0.05
PDE	33.46±0.39	0.58±0.05	70.52±4.71	94.69	-

5 Discussion

To better meet the demand for HR images in related fields, the study proposed a PDE ISRR method that introduced a priori information. Moreover, the experiments were conducted on the Flickr2K dataset and DIV2K dataset. The results revealed that in the Flickr2K dataset, the GME was the smallest when the GI of the image was extracted using CNN when compared with other methods. When the number of iterations was 500, the GME was 0.34. The GOC was at its maximum when the GI of the image was extracted using a CNN. When the number of iterations was 500, the GOC at this point was 0.94. The fusion a priori PDE model performed well when SRR was applied to the image. At an iteration of 500, the EEI and SDR were 0.64 and 0.05, respectively. The PSNR and SSIM were 41.74 dB and 0.97, respectively. The PSNR of bicubic interpolation was only 33.8dB. In the DIV2K dataset, the average PSNR of the PDE model that integrated prior information was 33.46dB, significantly higher than the 30.24dB of SRGAN.

The experimental results indicated that the proposed ISRR method provided richer guidance information for the PDE model by introducing prior information. This made the reconstructed image more accurate in terms of detail restoration and edge preservation. Second, the ResNet model was improved by removing the BNL and introducing a CAM and a gradient-guided branch. This effectively solved the problem of gradient dispersion or explosion while enhancing the model's ability to perceive key areas in an image. However, while the proposed model improved the accuracy of ISRR, it also came with higher computational costs. The CNN model contained a large number of parameters and layers, resulting in a complex computational process that required high computational resources and time costs. In practical applications, a balance needed to be struck between specific needs and resource constraints. Therefore, future research should further explore model compression techniques, such as pruning and quantization, to reduce parameter and computational complexity while maintaining model performance.

6 Conclusion

HR photographs are frequently compressed into LR images throughout the image transmission and storage process in an effort to increase efficiency and conserve space, which results in the loss of image features. However, the ISRR technique can recover more detail information from LR images. The suggested approach produces superior results on a number of common datasets and successfully raises the quality and resilience of ISRR. However, CNNs have a large number of layers and parameters, which require a large amount of computational time and storage space, and thus have a high demand on computational resources. Subsequent studies will try to use other methods to improve it and avoid the long computation time.

References

[1] Shaowei Zhang, Rongwang Yin, and Mengzi Zhang. Dynamic unstructured pruning neural network image super-resolution reconstruction. Informatica, 48(7):11-22, 2024. https://doi.org/10.31449/inf.v48i7.5332

[2] Chitwan Saharia, Jonathan Ho, William Chan, Tim Salimans, David J. Fleet, and Mohammad Norouzi. Image super-resolution via iterative refinement. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(4):4713-4726, 2022. https://doi.org/10.48550/arXiv.2104.07636

[3] Xiaoyan Wang, and Ya Li. Edge detection and simulation analysis of multimedia images based on intelligent monitoring robot. Informatica, 48(5):97-109, 2024. https://doi.org/10.31449/inf.v48i5.5366

[4] Xu Yan. A face recognition method for sports video based on feature fusion and residual recurrent neural network. Informatica, 48(12):137-152, 2024. https://doi.org/10.31449/inf.v48i12.5968

[5] Liqun Shan, Xueyuan Bai, Chengqian Liu, Yin Feng, Yanchang Liu, and Yanyan Qi. Super-resolution reconstruction of digital rock CT images based on residual attention mechanism. Advances in Geo-Energy Research, 6(2):157-168, 2022. https://doi.org/10.46690/ager.2022.02.07

[6] Kai Fukami, Koji Fukagata, and Kunihiko Taira. Super-resolution analysis via machine learning: A survey for fluid flows. Theoretical and Computational Fluid Dynamics, 37(4):421-444, 2023. https://doi.org/10.1007/s00162-023-00663-0

[7] Shuang Cong, and Yang Zhou. A review of convolutional neural network architectures and their optimizations. Artificial Intelligence Review, 56(3):1905-1969, 2023. https://doi.org/10.1007/s10462-022-10213-5

[8] Hamam Mokayed, Tee Zhen Quan, Lama Alkhaled, and V. Sivakumar. Real-time human detection and counting system using deep learning computer vision techniques. Artificial Intelligence and Applications. 1(4):221-229, 2023. https://doi.org/10.47852/bonviewAIA2202391

[9] Musa Alshawabkeh, Mohammad Hashem Ryalat, Osama M. Dorgham, Khalid Alkharabsheh, Mohammad Hjouj Btoush, and Mamoun Alazab. A hybrid convolutional neural network model for detection of diabetic retinopathy. International Journal of Computer Applications in Technology, 70(3-4):179-196, 2022. https://doi.org/10.1504/ijcat.2022.130886

[10] Xiewei Xiong, Tong Zhu, Yun Zhu, Mengyao Cao, Jin Xiao, Li Li, Fei Wang, Chunhai Fan, and Hao Pei. Molecular convolutional neural networks with DNA regulatory circuits. Nature Machine Intelligence, 4(7):625-635, 2022. https://doi.org/10.1038/s42256-022-00502-7

[11] Harsh Gupta, Naman Bansal, Swati Garg, Hritesh Mallik, Anju Prabha, and Jyoti Yadav. A hybrid convolutional neural network model to detect COVID-19 and pneumonia using chest X-ray images. International Journal of Imaging Systems and Technology, 33(1):39-52, 2023. https://doi.org/10.1002/ima.22829

[12] Yüksel Çelik, Erdal Başaran, and Yusuf Dilay. Identification of durum wheat grains by using hybrid convolution neural network and deep features. Signal, Image and Video Processing, 16(4):1135-1142, 2022. https://doi.org/10.1007/s11760-021-02094-y

[13] Yesim Eroglu, Muhammed Yildirim, and Ahmet Cinar. mRMR-based hybrid convolutional neural network model for classification of Alzheimer's disease on brain magnetic resonance images. International Journal of Imaging Systems and Technology, 32(2):517-527, 2022. https://doi.org/10.1002/ima.22632

[14] Youngmin Jeon, and Donghyun You. Super-resolution reconstruction of transitional boundary layers using a deep neural network. International Journal of Aeronautical and Space Sciences, 24(4):1015-1031, 2023. https://doi.org/10.1007/s42405-023-00598-0

[15] Weizhi Du, and Shihao Tian. Transformer and GAN-based super-resolution reconstruction network for medical images. Tsinghua Science and Technology, 29(1):197-206, 2023. https://doi.org/10.26599/TST.2022.9010071

[16] Yao Sui, Onur Afacan, Camilo Jaimes, Ali Gholipour, and Simon K Warfield. Scan-specific generative neural network for MRI super-resolution reconstruction. IEEE Transactions on Medical Imaging, 41(6):1383-1399, 2022. https://doi.org/10.1109/TMI.2022.3142610

[17] Mingjin Zhang, Qianqian Wu, Jie Guo, Yunsong Li, and Xinbo Gao. Heat transfer-inspired network for image super-resolution reconstruction. IEEE Transactions on Neural Networks and Learning Systems, 35(2):1810-1820, 2022. https://doi.org/10.1109/TNNLS.2022.3185529

[18] Bo Zhang, Wei Xiong, Muyuan Ma, Mingqing Wang, Dong Wang, Xing Huang, Le Yu, Qiang Zhang, Hui Lu, Danfeng Hong, Fan Yu, Zidong Wang, Jie Wang, Xuelong Li, Peng Gong, and Xiaomeng Huang. Super-resolution reconstruction of a 3 arc-second global DEM dataset. Science Bulletin, 67(24):2526-2530, 2022. https://doi.org/10.1016/j.scib.2022.11.021

[19] Jingmin Gong, Qinfei Xu, and Qinfei Xu. Temporal transformer-based video super-resolution reconstruction with cross-modal attention. Informatica, 49(10):179-190, 2025. https://doi.org/10.31449/inf.v49i10.7146

[20] Saeed Iqbal, Adnan N. Qureshi, Jianqiang Li, and Tariq Mahmood. On the analyses of medical images using traditional machine learning techniques and convolutional neural networks. Archives of Computational Methods in Engineering, 30(5):3173-3233, 2023. https://doi.org/10.1007/s11831-023-09899-9

[21] Fanghui Chen, Shouliang Li, Jiale Han, and Fengyuan Ren. Review of lightweight deep convolutional neural networks. Archives of Computational Methods in Engineering, 31(4):1915-1937, 2024. https://doi.org/10.1007/s11831-023-10032-z

[22] D. R. Sarvamangala, and Raghavendra V. Kulkarni. Convolutional neural networks in medical image understanding: A survey. Evolutionary Intelligence, 15(1):1-22, 2022. https://doi.org/10.1007/s12065-020-00540-3

[23] Andi Johnson. Investigation of network models finite difference method. Eurasian Journal of Chemical, Medicinal and Petroleum Research, 2(1):1-9, 2023. https://doi.org/10.5281/zenodo.7347257

[24] Hao-Jun Michael Shi, Melody Qiming Xuan, Figen Oztoprak, and Jorge Nocedal. On the numerical performance of finite-difference-based methods for derivative-free optimization. Optimization Methods and Software, 38(2):289-311, 2023. https://doi.org/10.48550/arXiv.2102.09762

[25] Ram Shiromani, Vembu Shanthi, and J. Vigo-Aguiar. A finite difference method for a singularly perturbed 2-D elliptic convection-diffusion PDEs on Shishkin-type meshes with non-smooth convection and source terms. Mathematical Methods in the Applied Sciences, 46(5):5915-5936, 2023. https://doi.org/10.1002/mma.8877

[26] Daisuke Inoue, Yuji Ito, Takahito Kashiwabara, Norikazu Saito, and Hiroaki Yoshida. A fictitious-play finite-difference method for linearly solvable mean field games. ESAIM: Mathematical Modelling and Numerical Analysis, 57(4):1863-1892, 2023. https://doi.org/10.1051/m2an/2023026