ISSN 1330-3651 (Print), ISSN 1848-6339 (Online) https://doi.org/10.17559/TV-20230717000808
Original scientific paper

RoughSet-DDPM: An Image Super-Resolution Method Based on Rough set Denoising Diffusion Probability Model

Tao SONG*, Ran WEN, Lei ZHANG

Abstract: Image super-resolution aims to generate high-resolution (HR) images from low-resolution (LR) inputs. Existing methods like autoregressive models, generative adversarial networks (GANs), and denoising diffusion probability models (DDPMs) have limitations in image quality or sampling efficiency. This paper proposes Rough Set-DDPM, a new super-resolution technique combining rough set theory and DDPMs. The rough set formulation divides the DDPM sampling sequence into optimal sub-columns by minimizing roughness of sample sets. Particle swarm optimization identifies the sub-columns with lowest roughness. Rough Set-DDPM applies iterative denoising on these optimal columns to output HR images. Experiments on the FFHQ dataset validate that Rough Set-DDPM improves DDPM sampling efficiency while maintaining image fidelity. Quantitative results show Rough Set-DDPM requires fewer sampling steps and generates higher quality HR images compared to autoregressive models and GANs. By enhancing DDPM sampling, Rough Set-DDPM provides an effective approach to super-resolution that balances image quality and sampling speed. The key contributions include introducing rough sets to optimize DDPM sampling and demonstrating superior performance over existing methods.

Keywords: image super-resolution; probability model for denoising diffusion; rough set; U-net network

1 INTRODUCTION

Super-resolution (SR) images are obtained by converting low-resolution (LR) images into high-resolution (HR) images. SR images are useful in various professional fields, including medical imaging, security monitoring, and remote sensing satellites. The depth generation model, owing to its ability to capture diverse examples and intricate image experience distributions, is extensively applied in super-resolution imaging. Remarkable results have been achieved by depth generation models in learning complex image experience distributions [1-3]. Methods such as autoregressive models (AR) [4-6], variational autoencoders (VAEs) [7-8], normalized flow (NFs) [9-10], generative adversarial networks (GANs) [11-13], and denoising diffusion probability models (DDPM) have demonstrated satisfactory image generation capabilities. These models have also been applied to SR imaging in conditional generation tasks. Autoregressive models [14-15] have been successfully employed in SR and cascading upsampling. However, their high computational costs limit their applicability to LR images. While the accurate modeling of normalized flow [14-16] improves sampling speed, its expressive ability is constrained by the requirement to process the reversible parametric transformation of the Jacobian determinant. Although VAEs [17-19] offer faster sampling, the quality of image generation is lower compared with GANs and NFs. GANs provide a super-resolution method based on conditional image generation. Cascaded GANs [20-22] have achieved significant progress in generating HR images, with FSRGAN demonstrating high-quality super-resolution of faces [23-24]. However, many GAN-based methods often face challenges in optimization because of unstable generator and discriminator training. Therefore, auxiliary objective functions are necessary to ensure consistency with low-resolution inputs. The unified loss function design of DDPM [25] exhibits simplicity and provides sufficient theoretical support for stable training. However, DDPM has an inherent disadvantage compared with generative adversarial networks, autoencoders, and normalized flows-namely, the excessive number of sampling steps, resulting in longer sampling times. However, improving sampling efficiency while maintaining image quality in DDPM remains an important challenge. The existing super-resolution methods include the High Resolution Image Synthesis with Latent Diffusion Models method proposed by Robin Rombach and Andreas Blattmann et al., which utilizes diffusion models in potential space to improve image resolution. Although it further improves computational power, it can cause information loss in image details during the encoding and decoding process. Secondly, the efficiency of image sampling was not effectively improved during the sampling process; the Image Super Resolution via Iterative Refinement method proposed by Chitwan Saharia and Jonathan Ho et al. conditionally deepens the cyclic iterative denoising sampling on the basis of DDPM. Although it can obtain super-resolution images, DDPM has high GPU consumption and computational power during image space training, and secondly, the efficiency of cyclic iterative denoising sampling is very low. The Taming Transformers for High Resolution lmage Synthesis method proposed by Patrick Esser and Robin Rombach et al., with the contribution of VQ-GAN, obtains stunning 1280 × 640 ultra-high pixel natural landscape images, which consume less GPU and computational power than the first two methods. However, the process of vector quantization may lose some information on image details and the limitations of GAN's own convergence may to some extent affect the acquisition of ultra-high pixel images. The above are the mainstream methods for obtaining super-resolution images, all based on the generation model for obtaining super-resolution images. The method proposed in this article can compensate for the loss of detail information in the encoding and decoding images in the potential space, and effectively solve the problem of low sampling efficiency in DDPM by using the sampling sequence obtained by rough set method. However, this method consumes some GPU during training and has high computational power. For this problem, this paper uses high-performance GPU to train the model. Rough set theory is a new mathematical tool that characterizes imprecise, uncertain, incomplete, and

inconsistent information. It can analyze, infer, and mine implicit knowledge and rules from a large amount of data. Neural networks have excellent numerical approximation and generalization abilities, which can process quantitative and numerical information. Compared to rough set data processing, they can obtain more precise results. However, when the network size is large and there are many samples, the training process becomes complex and diffuse, which limits the promotion of neural network usability [26]. Combining rough set and neural network, using rough set to preprocess input information, that is, selecting the training set. In general, the training set often has a lot of redundancy, and using such a training set in neural networks often leads to overfitting. Rough set analysis can filter redundant information, thereby improving the generalization ability of neural networks [27-29]. The effective combination of rough sets and deep neural networks lays the theoretical foundation for improving the sampling efficiency of U-net trained data to rough sets in this paper. Based on this, the theory of rough sets is used to solve effective sampling sequences to improve sampling efficiency. Compared with classical rough sets, variable precision rough sets, and fuzzy rough sets, RoughSet-DDPM not only possesses the characteristics of rough sets, but also effectively utilizes it in multimodal fields such as image processing, text generation, and video generation. When using rough set theory to solve effective sampling sequence equations, convergence is slow. To improve the efficiency of solving effective subsequences, particle swarm optimization algorithm is used to optimize the process of solving effective subsequences and improve the efficiency of solving effective subsequences. This study established a rough set denoising diffusion probability method to overcome the limitations of deep generative models in image super-resolution. This method improves upon the Markov chain-free diffusion probability model (DDIM) proposed by Nichol A[30] et al., who demonstrated that the diffusion process of DDPM does not rely on Markov chains. Although selecting random sub-columns improves sampling efficiency, it does not ensure the robustness of the sampled images. In this study, we combined rough set theory with the particle swarm optimization algorithm to obtain better sub-columns, ensuring the robustness of the sampled images. Simultaneously, the SR image is obtained by applying an in-depth iterative denoising process to the sub-columns.

2 THEORETICAL BASIS

2.1 Rough set

Given a knowledge base $K=(U,R)$, for each subset $X \subseteq U$ and an equivalence relation R, two subsets are defined:

$$\underline{R}X = \cup\{Y \in U/R \mid Y \subseteq X\} \tag{1}$$

$$\overline{R}X = \cup\{Y \in U/R \mid Y \cap X \neq \varnothing\} \tag{2}$$

Call the R lower approximation set and the upper approximation set of X, respectively. The approximate roughness of the set X defined by the equivalence relation R can be expressed as follows:

$$\rho_R(X) = 1 - \frac{|\underline{R}X|}{|\overline{R}X|} \tag{3}$$

2.2 Particle Swarm

Particle swarm optimization (PSO) employs the individual optimal solution to search for the global optimal solution. The algorithm utilizes the speed and position update formula with weight, which is presented as follows.

$$v_i^{k+1} = \omega v_i^k + c_1 rand_1\left(pbest_i^k - x_i^k\right) + \\ +c_2 rand_2\left(gbest^k - x_i^k\right) \tag{4}$$

$$x_i^{k+1} = x_i^k + v_i^k \tag{5}$$

2.3 DDIM

The DDIM is a development based on DDPM. In DDPM, both the forward diffusion and backward diffusion processes rely on Markov chains. The resulting loss function, denoted as L simple, represents the marginal distribution of dependencies and is independent of the joint distribution. The non-Markov chain-based denoising diffusion probability model (DDIM), as proposed in the literature [30], utilizes the same L simple as DDPM. The diffusion process in DDIM is divided into forward diffusion and backward diffusion, both of which are independent of the Markov chain. Fig. 1 illustrates the DDIM diffusion process, while Fig. 2 showcases the DDPM diffusion process. DDIM introduces a hyperparameter to the non-Markov chain forward diffusion process $\sigma \in \mathrm{R}_{\geq 0}^{T}$, and the joint distribution of the forward diffusion is set as follows:

$$q_\sigma(x_{1:T} \mid x_0) := q_\sigma(x_T \mid x_0) \prod_{t=2}^{T} q_\sigma(x_{t-1} \mid x_t, x_0) \tag{6}$$

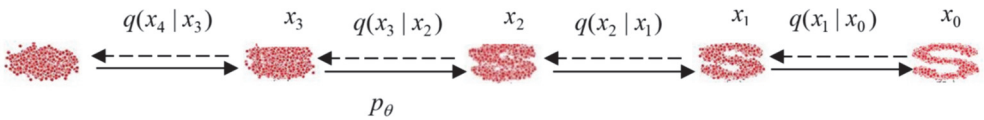


Figure 1 Diffusion process of DDIM

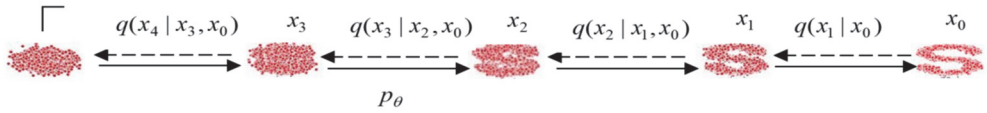


Figure 2 Diffusion process of DDPM

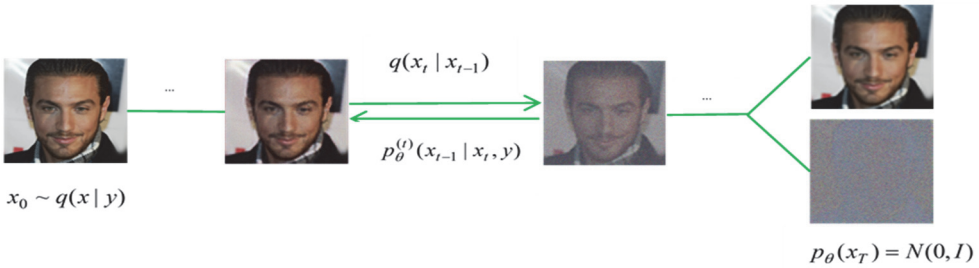


Figure 3 Conditional DDIM diffusion process

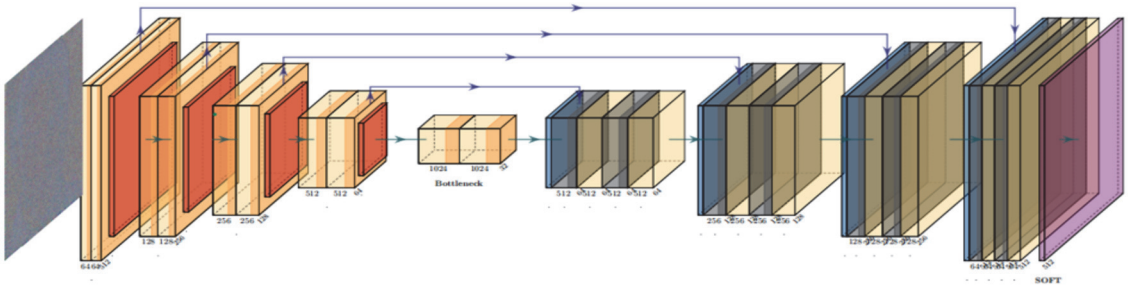


Figure 4 U-net network framework

Among them, $q_\sigma(x_T \mid x_0) = N(\sqrt{\alpha_T} x_0, (1-\alpha_T)I)$ $(t>1)$ this marginal distribution is the expectation of $\sqrt{\alpha_T} x_0$, $(1-\alpha_T)I$ is the variance. The posterior distribution at time T is expressed as follows :

$$q_\sigma(x_{t-1} \mid x_t, x_0) = N\left(\sqrt{\alpha_{t-1}} x_0 + \sqrt{1-\alpha_{t-1}-\sigma_t^2} \cdot \frac{x_t - \sqrt{\alpha_t} x_0}{\sqrt{1-\alpha_t}}, \sigma_t^2 I\right) \quad (7)$$

The DDIM had a marginal distribution of $q_\sigma(x_t \mid x_0) = N(\sqrt{\alpha_t} x_0, (1-\alpha_t)I)$.

This model encompasses the design of a non-Markov chain forward diffusion process, with its edge distribution being consistent with DDPM. Therefore, DDIM trains the model using the same Lsimple as DDPM.

$$x_t = \sqrt{\alpha_t} x_0 + \sqrt{(1-\alpha_t)}\varepsilon, (\varepsilon \sim N(0, I)) \quad (8)$$

DDIM employs a non-Markov chain inverse diffusion process, which serves as a denoising procedure to approximate the posterior distribution of the forward diffusion process at any given time. A trainable inverse generation process is defined, and at each iteration, an approximation of the process is obtained. The variance at time T remains constant, and expectations are involved. The U-net network is utilized. This approximation is then incorporated into Eq. (8) to predict the expected value and facilitate the denoising approximation of the process.

Substituting x_t and $\varepsilon_\theta^{(t)}(x_t)$ into Eq. (8), we can obtain that x_0 at time t is the denoised observation quantity denoted by $f_\theta^{(t)}(x_t)$.

$$f_\theta^{(t)}(x_t) := \left(x_t - \sqrt{1-\alpha_t} \cdot \varepsilon_\theta^{(t)}(x_t)\right) \Big/ \sqrt{\alpha_t} \quad (9)$$

Inverse diffusion of DDIM is the process from time to 1, the $p_\theta(x_T) = N(0, I)$ is in the process of the Gaussian noise, at any time for the spread of the process:

$$p_\theta^{(t)}(x_{t-1} \mid x_t) = \begin{cases} N(f_\theta^{(1)}(x_1), \sigma_1^2 I), t = 1 \\ q_\sigma(x_{t-1} \mid x_t, f_\theta^{(t)}(x_t)), t > 1 \end{cases} \quad (10)$$

When $t>1$, the inverse diffusion x_{t-1} is obtained from Eq. (7), Eq. (9), and Eq. (10).

$$x_{t-1} = \sqrt{\alpha_{t-1}}\left(\frac{x_t - \sqrt{1-\alpha_t}\varepsilon_\theta^{(t)}(x_t)}{\sqrt{\alpha_t}}\right) + \sqrt{1-\alpha_{t-1}-\sigma_t} \cdot \varepsilon_\theta^{(t)}(x_t) + \sigma_t \varepsilon_t \quad (11)$$

In reference [21], when $\sigma_t = 0$ is the DDIM model, $\sigma_t = 1$ represents the DDPM model. According to the literature [21], $L_{simple} = \| \varepsilon(x_t) - \varepsilon_\theta^{(t)}(\sqrt{\alpha_t} x_0 + \sqrt{1-\alpha_t}\varepsilon) \|^2$ is consistent with the loss function of DDPM. The U-net network is trained with the Lsimple loss function for DDIM.

3 METHOD

When DDIM is utilized to denoise $p_\theta(x_T)$ from full Gaussian noise step by step, the result is a high-resolution image x_0. In this process, a gap between the high-resolution image is obtained during each deep iteration denoising step, and y may differ from the input low-resolution image. To ensure that the same information as the input LR

images x_0 and y is generated, this paper proposes the application of conditional diffusion DDIM to enhance low-resolution images and guarantee the consistency of information between x_0 and y. According to Eq. (10), the inverse diffusion process of conditional DDIM is obtained. The diffusion process of the model is shown in Fig. 3. The loss function used in this method is the same as that of DDPM and DDIM in the inverse diffusion process. It is independent of the Markov chain in the diffusion process. The rough set theory is employed to obtain a more optimal subsequence. Each instance in the diffusion process corresponds to an element in U. From time T to 1, each corresponding element at time t forms the universe of discourse in the rough set. The loss function is derived as follows: the most critical step to generate the diffusion model is the construction of the generation process $p(x_{t-1} \mid x_t)$, and for the generation with y as the input condition is $p(x_{t-1} \mid x_t)$ replaced $p(x_{t-1} \mid x_t, y)$, and the input y is added in the generation process. Using Bayes' theorem.

$$p(x_{t-1} \mid y) = \frac{p(x_{t-1}) p(y \mid x_{t-1})}{p(y)} \tag{12}$$

In each complement condition x_t, get:

$$p(x_{t-1} \mid x_t, y) = \frac{p(x_{t-1} \mid x_t) p(y \mid x_{t-1}, x_t)}{p(y \mid x_t)} = \\ = p(x_{t-1} \mid x_t) \mathrm{e}^{\log p(y \mid x_{t-1}) - \log p(y \mid x_t)} \tag{13}$$

Taylor in Eq. (13) $\log p(y \mid x_{t-1}) - \log p(y \mid x_t)$ Spread out with Taylor.

$$\log p(y \mid x_{t-1}) - \log p(y \mid x_t) \approx (x_{t-1} - x_t) \times \nabla_{x_t} \log p(y \mid x_t) \tag{14}$$

Available by Eq. (14)

$$p(x_{t-1} \mid x_t, y) \propto e^{-\|x_{t-1} - \mu(x_t)\|^2 / 2\sigma_t^2 + (x_{t-1} - x_t) \nabla_{x_t} \log(y \mid x_t)} \\ \propto e^{-\|x_{t-1} - \mu(x_t) - \sigma_t^2 \nabla_{xt} \log p(y \mid x_t)\|^2 / 2\sigma_t^2} \tag{15}$$

According to the Eq. (15) and $p(x_{t-1} \mid x_t, y) = \mathrm{N}(x_{t-1}; \mu(x_t, y), \sigma_t^2 I)$ get.

$$\mu(x_t, y) = \frac{1}{\alpha_t} \left(x_t - \frac{\beta_t^2}{\overline{\beta}_t} \varepsilon_\theta(x_t, y, t) \right) \tag{16}$$

The loss function trained from the above formula is:

$$\mathrm{E}_{x_0, y \sim p(x_0, y), \varepsilon \sim N(0, I)} [\| \varepsilon - \varepsilon_\theta(\sqrt{\alpha_t} x_0 + \sqrt{1 - \alpha_t} \varepsilon, y, t) \|^2] \tag{17}$$

The Eq. (17) is further simplified to obtain the final training loss function.

$$L'_{simple} = \| \varepsilon(x_t) - \varepsilon_\theta^{(t)}(\sqrt{\alpha_t} x_0 + \sqrt{1 - \alpha_t} \varepsilon, y) \|^2 \tag{18}$$

The U-net network is trained using the unified objective Eq. (18) to establish the conditional diffusion DDIM model. In this paper, the framework design of U-net network in the inverse diffusion denoising process of this model is shown in Fig. 4. DDIM selects a subsequence randomly from the original sampling sequence, which improves sampling efficiency. However, this random selection does not guarantee better sampling results compared to DDPM each time. To this end, this study adopted a rough set denoising diffusion probability method based on DDIM, which aims to improve the selection of sub-columns. The subscript of an element in U represents the distribution of x at a specific time t, while the subscripts represent the distribution of t at different time points. The diffusion ∇t on process is non-Markovian, and the distribution corresponding to each time t is independent x_t. A time period can be seen as an equivalence relation R. The final determination of the knowledge base $K = (U, \nabla t)$, ∇t is related to the size of T. The equivalence class obtained by partitioning the universe U using the equivalence relation ∇t is denoted as $U/\nabla t$. Based on the definition in 1.1, the lower and upper approximation sets and roughness of any sampled column $Q(\mathrm{Q} \subseteq \mathrm{U})$ on the domain U are determined:

Lower approximation set: $\underline{\nabla t} Q = \cup \{ Y \in U/\nabla t \mid Y \subseteq Q \}$.

Coapproximation set： $\overline{\nabla t} Q = \cup \{ Y \in U/\nabla t \mid Y \cap Q \neq \varnothing \}$.

Degree of roughness:

$$\rho_{\nabla t}(Q) = 1 - \frac{|\underline{\nabla t} Q|}{|\overline{\nabla t} Q|} \tag{19}$$

Based on the principles of rough set theory [31], set imprecision arises from the presence of a boundary region. The roughness of a set decreases as the boundary region becomes smaller, indicating a higher degree of certainty in the knowledge. In the universe U, when the roughness of a subset Q is minimized, the sequence formed by the subscripts of each element in Q represents a more optimal sample sequence. However, due to the large size of the domain U, there exist numerous subsets. The particle swarm optimization algorithm is employed to address the challenge of efficiently obtaining improved image sequences within the domain U. This algorithm utilizes roughness as the objective function to solve the problem at hand. The specific process of the method is as follows:

U-net trains the conditional diffusion DDIM model with L'_{simple} as the loss function.

1：The probability distribution $q(x_0)$ of x_0.

2：$t \in \{1, 2, \ldots, T-1, T\}$

3：ε is taken randomly from the standard normal distribution $N(0, \mathrm{I})$.

4：Deep learning network training.

$L'_{simple} = \| \varepsilon(x_t) - \varepsilon_\theta^{(t)}(\sqrt{\alpha_t} x_0 + \sqrt{1 - \alpha_t} \varepsilon, y) \|^2$

5：Completion of training.

Sampling

1 : $x_T \sim N(0, I)$

2 : Initialize the particle position in Eq, (1) Eq. (4), and $c_1 = c_2 = 2$, $rand_1$ 和 $rand_2$ take a random number in $x[0, 1]$. The inertia weight is $\omega = 0.6$

3 : for $i = 1, ..., 2^{|U|}$

4 : $Q_i \subseteq U$, $\rho_{\nabla t}(Q_i)$

5 : if $(\rho_{\nabla t}(Q_i) < pbest_i)$

6 : $pbest_i = \rho_{\nabla t}(Q_i)$

7 : if $(pbest_i < gbest)$

8 : $gbest = pbest_i$

9 : Evolve the particle positions and velocities in Eq. (4) and Eq. (5).

10 : break

11 : end for

12 : The index sequence $\{t_i, t_{i+1}, .., t_{i+j}\}$ of the elements of the set Q_i, when $\rho_{\nabla t}(Q_i)$ is minimized, is the best subsequence $j = |Q_i|$ to sample.

13 : for $t = t_i, t_{i+1}, ..., t_{i+j}$

14 : $Z \sim (0, I)$

15 : if $(t > t_{i+j})$

16 : Eq. (8)

$$x_{t-1} = \sqrt{\alpha_{t-1}}\left(\frac{x_t - \sqrt{1-\alpha_t}\varepsilon_\theta^{(t)}(x_t)}{\sqrt{\alpha_t}}\right) + \sqrt{1-\alpha_{t-1}} \cdot \varepsilon_\theta^{(t)}(x_t)$$

17 : end for

18 : return x_0

4 ANALYSIS OF EXPERIMENT

This paper conducts a comparative analysis using the FFHQ dataset, utilizing the Pytorch deep learning framework (GPU version) and an RTX A6000 graphics card. Based on the selected dataset and hyperparameter $\alpha_t(\{\alpha_t \in (0,1)\}_{t=1}^{T})$, T = 1000 was determined. Fig. 5 illustrates a portion of the training process of the proposed model on the FFHQ dataset.

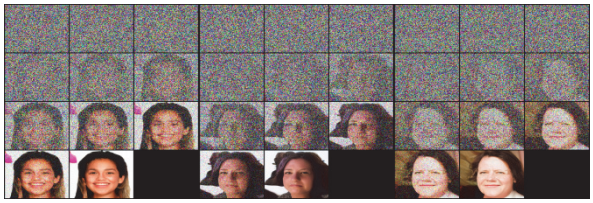


Figure 5 Training process diagram

When T = 1000, Eq. (13) serves as the objective function for the domain $U = \{x_{1000}, x_{999}, x_{998}, ..., x_t, ..., x_2, x_1\}$. By utilizing the particle swarm optimization algorithm, the minimum roughness is obtained, resulting in a more optimal sequence of subscripts within the domain. The sequence of improved sampled looks obtained is denoted as Q. Fig. 6 presents a comparison of the sampling results on Q of the proposed method, the sampling results of DDPM on the original sequence, and the sampling results of DDIM on three randomly selected subcolumns of equal length.

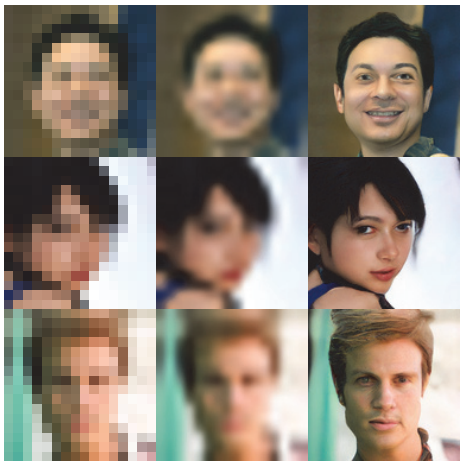
(a) DDIM subcolumn1 (b) DDIM sub column 2, (c) DDIM subcolumn3, (d) DDPM, (e) Method proposed in this study

Figure 6 Comparison of sampling results for different subcolumns

Table 1 Fig. 6 FID values of the sampling results

FID			
Fig. 6a DDIMSubcolumn 1	329.99	338.51	309.16
Fig. 6b DDIMSubcolumn 2	227.23	221.69	219.91
Fig. 6c DDIMSubcolumn 3	105.79	107.25	90.67
Fig. 6d DDPM Original sequence	21.98	22.52	21.02
Fig. 6e Method of this paper	22.06	22.67	21.39

Fig. 6 and Table 1 clearly demonstrate that the sampling effect of DDIM in three random subsequences surpasses that of DDPM and the proposed method in terms of sequence results. These findings indicate that the sampling sequence obtained in this study not only achieves effective sampling but also yields superior sampling outcomes. The application of this method on the face dataset FFHQ covers a range from x LR to y SR, as illustrated in Fig. 7 and Fig. 8. Fig. 7 provides a visual representation of the different stages involved in the proposed method. LR represents the low-resolution image, INF represents the intermediate image obtained during the process from low-resolution to high-resolution, and SR represents the final super-resolution image. It showcases the LR $16\times16\rightarrow128\times128$ SR images obtained through deep iterative denoising using the proposed method. On the other hand, Fig. 8 presents the super-resolution sampling process obtained through the proposed method.



LR INF SR

Figure 7 LR $16\times16\rightarrow128\times128$ SR

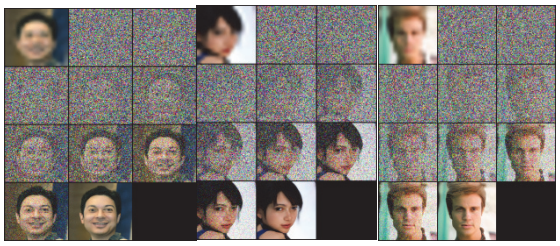


Figure 8 128×128SR Sampling process

LR INF SR

Figure 9 LR $64\times64\rightarrow512\times512$ SR

The application of the proposed method from 64×64 low-resolution to 512×512 super-resolution on the face dataset FFHQ is illustrated in Fig. 9 and Fig. 10.

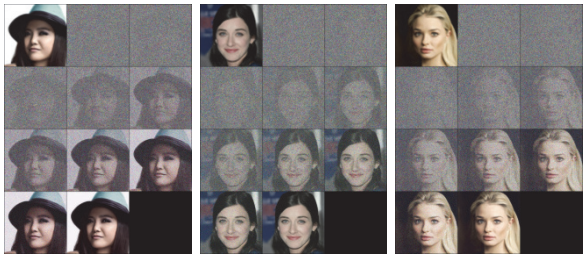
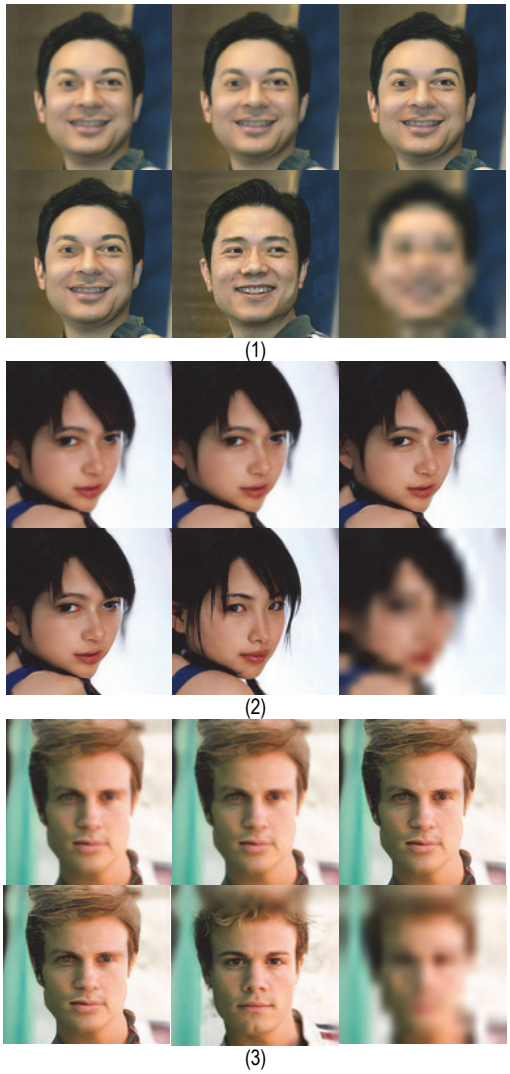


Figure 10 SR sampling process

In Fig. 9, LR represents the low-resolution image, INF represents the intermediate image obtained during the process from low-resolution to high-resolution, and SR represents the final super-resolution image. Fig. 7 displays the LR and SR images obtained through deep iterative denoising using the proposed method. Fig. 10, on the other hand, illustrates the SR sampling process obtained through the proposed method.



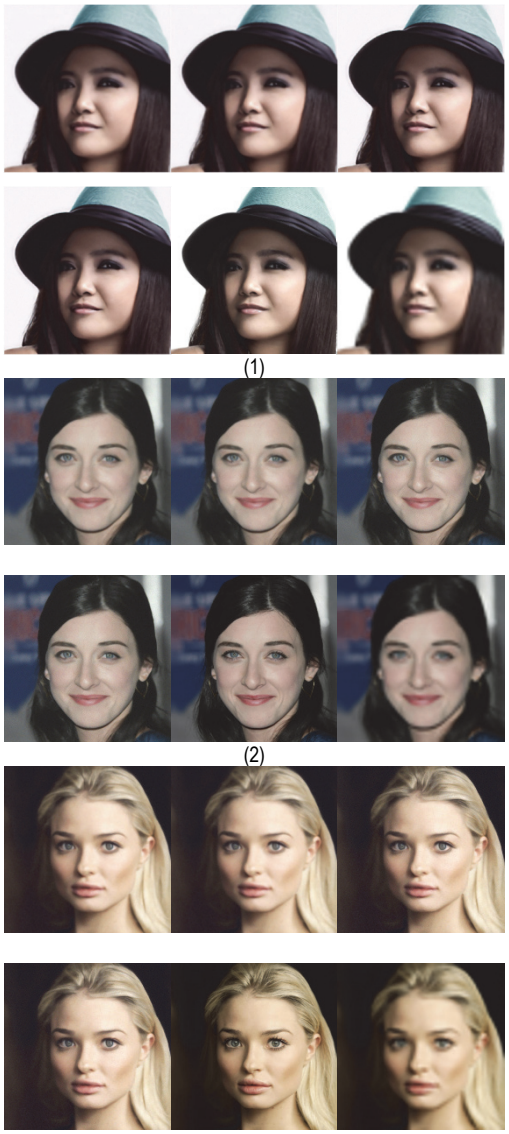
(1) (2) (3)

(a) AR, (b) GAN, (c) DDPM, (d) Method of this paper, (e) HR, (f) INF

Figure 11 SR of the four methods

Furthermore, Fig. 11 and Fig. 12 showcase the SR and SR images acquired through the autoregressive model,

adversarial neural network, denoising diffusion probabilistic model, and the proposed method on the FFHQ dataset.



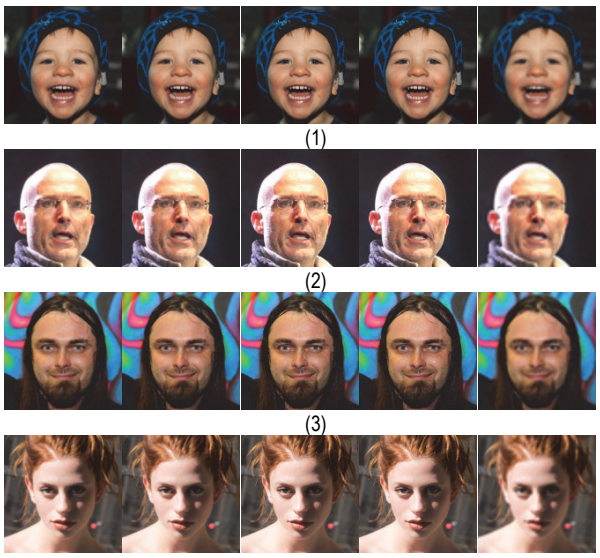
(a) AR, (b) GAN, (c) DDPM, (d) Method proposed in this paper, (e) HR, (f) INF

Figure 12 SR of the four methods

In this study, FID, SSIM, and PSNR were used to evaluate the quality of SR images obtained by the four methods shown in Fig. 11 and Fig.12. The smaller the FID, the more similar the images, the larger the SSIM, the more similar the structure of two images, and the larger the PSNR, the less the image distortion. According to Tab. 2, the FID of d(1), d(2), and d(3) in Fig. 11 and Fig. 12 was smaller than that of a(1), b(1), a(2), b(2), and a(3), b(3), and the FID of b(1), b(2), and b(3) was smaller than that of a(1), a(2), and a(3), respectively. The PSNR and SSIM of d(1), d(2), and d(3) were greater than those of a(1), b(1), a(2), b(2), a(3), b(3), b(1), b(2), and b(3), while the PSNR and SSIM of b(1), b(2), and b(3) were greater than those of a(1), a(2), and a(3). The FID, SSIM, and PSNR of d(1), d(2), and d(3) were comparable to those of c(1), c(2), and c(3). This indicated that the SR image quality obtained by the proposed method was comparable to that of DDPM and better than that of GAN. The SR image quality of GAN was better than that of AR. It further showed that obtaining SR images using the proposed method could make up for the defects of obtaining SR images using GAN and AR methods and could generate higher-quality SR images.

Tab. 2 presents the FID, PSNR, and SSIM values obtained from comparing the SR images generated by the four methods with the original HR image. The comparison is conducted on both the LR and SR images.

The SR images obtained on the CELEBA dataset using autoregressive model, adversarial neural network, denoising diffusion probability model, and our method are shown in Fig. 13.



(a) AR, (b) GAN, (c) DDPM, (d) proposed method, (e) INF

Figure 13 Comparison of four methods 512×512SR

Although the quality of SR images obtained by the proposed method was equivalent to that of the SR images obtained by DDPM, the efficiency of obtaining SR images by the proposed method was significantly higher than that of the SR images obtained by DDPM. The sampling times of the two methods to obtain SR images in Fig. 11 and Fig. 12 are shown in Tab. 3. Tab. 3 clearly demonstrates that the sampling time for SR images obtained using the proposed method is significantly shorter than that of DDPM, indicating a higher sampling efficiency. According to Tab. 3, it can be seen that DDPM has significantly higher time consumption in sampling than RoughSetDDPM. The sampling efficiency of the two methods is mainly compared in terms of time consumption. From the perspective of time complexity, both DDPM and RoughSetDDPM have a time complexity of $O(n^2)$. However, the sampling efficiency of DDPM images at a step size of 1000 is significantly lower than that of RoughSetDDPM at the same latitude. Because of the complexity of time n due to different scales, time complexity $O(n^2)$ on DDPM n of 1000 and the n of 180 on RoughSet-DDPM. From n the scale, the time complexity of DDPM is higher than that of RoughSet-DDPM, which also shows that the sampling efficiency of RoughSet-DDPM is higher than that of DDPM.

Table 2 Quantitative evaluation of SR image quality

Figure			(a) AR	(b) GAN	(c)DDPM	(d) Method proposed in this paper
Fig. 11 $16\times16 \rightarrow 128\times128$ SR	FID	(1)	43.33	20.42	*18.13*	18.05
		(2)	23.01	15.29	*10.87*	10.62
		(3)	32.39	20.84	*14.10*	14.02
	SSIM	(1)	0.71	0.74	*0.77*	0.77
		(2)	0.64	0.66	*0.71*	0.71
		(3)	0.75	0.76	*0.81*	0.81
	PSNR	(1)	27.57	28.02	*28.86*	28.88
		(2)	26.31	28.75	*29.62*	29.7
		(3)	26.44	28.90	*29.74*	29.83
Fig. 12 $64\times64 \rightarrow 512\times512$ SR	FID	(1)	44.84	33.67	*26.43*	26.11
		(2)	41.82	10.39	*7.40*	7.03
		(3)	43.00	19.00	*12.40*	10.57
	SSIM	(1)	0.76	0.80	*0.85*	0.85
		(2)	0.72	0.78	*0.82*	0.82
		(3)	0.75	0.78	*0.83*	0.83
	PSNR	(1)	25.78	27.80	*28.33*	28.56
		(2)	26.10	28.73	*29.80*	29.96
		(3)	27.43	28.36	*29.14*	29.54

Table 3 Sampling time of the proposed method and DDPM to obtain SR images

figure		c DDPM (s)	d Method of this paper (s)
		Sampling sequence length: 1000	Sampling sequence length: 180
Fig. 11 $16\times16 \rightarrow 128\times128$ SR	(1)	103	20
	(2)	104	22
	(3)	102	25
Fig. 12 $64\times64 \rightarrow 512\times512$ SR	(1)	125	32
	(2)	131	29
	(3)	128	30

Acknowledgements

This work was supported by the National Nature Science Foundation of China under Award Number 71864035.

5 CONCLUSIONS

In summary, this paper presented RoughSet-DDPM, a new technique for single image super-resolution. RoughSet-DDPM incorporates rough set theory to optimize the sampling process of denoising diffusion probability models (DDPMs). The rough set formulation divides the DDPM sampling sequence into sub-columns and minimizes the roughness of each subset. Particle swarm optimization efficiently identifies the optimal sub-columns with lowest roughness. RoughSet-DDPM applies iterative denoising on only these sub-columns to generate high-resolution output images. Experiments on the FFHQ dataset validate that RoughSet-DDPM improves sampling efficiency compared to standard DDPM. Quantitative results also demonstrate that RoughSet-DDPM requires fewer sampling steps while achieving higher image quality than autoregressive models and GANs for super-resolution. By enhancing DDPM sampling through rough sets, RoughSet-DDPM balances sampling speed and image fidelity. The key contributions of this work include:

1) Introducing rough sets to optimize DDPM sampling sequences for super-resolution.
2) Employing particle swarm optimization to identify optimal sub-columns.
3) Demonstrating improved performance over existing super-resolution methods.

Future work involves reducing the computational complexity of RoughSet-DDPM for large-scale datasets. Investigating extensions to other image and video generation tasks is another promising direction. The sampling efficiency and image quality improvements demonstrated by RoughSet-DDPM suggest it could enable new applications in computational photography.

6 REFERENCES

[1] Sutskever, I., Vinyals, O., & Le Quoc, V. (2014). Sequence to Sequence Learning with Neural Networks. *CoRR*, *2*, 3104-3112.

[2] Petrović, A., Biserčić, A., Delibašić, B., & Milenković, D. (2022). Machine Learning approach for learning temporal point process. *Computer Science and Information Systems*, *19*(2), 1007-1022. https://doi.org/10.2298/CSIS210609016P

[3] Vaswani, A., Shazeer, N., & Parmar N., et al. (2017). Attention Is All You Need. *CoRR*, 6000-6010.

[4] Oord, D. V. A., Dieleman, S., & Zen, H., et al. (2016). WaveNet: A Generative Model for Raw Audio. *CoRR*, 5210-5216.

[5] Teršek, M., Kljun, M., Peer, P., & Emeršič, Ž. (2022). Re-evaluation of the CNN-based state-of-the-art crowd-counting methods with enhancements. *Computer Science and Information Systems*, *19*(3), 1177-1198. https://doi.org/10.2298/CSIS210925023T

[6] Van den Oord, A., Kalchbrenner, N., Vinyals, O., Espeholt, L., Graves, A., & Kavukcuoglu, K. (2016). Conditional Image Generation with Pixel CNNDecoders. *CoRR*, 4797-4805.

[7] Giordano, R., Broderick, T., & Jordan, M. I. (2018). Covariances, Robustness, and Variational Bayes. *Journal of Machine Learning Research*, *19*(1), 1981-2029.

[8] Cornillère, V., Djelouah, A., & Yifan, W., et al. (2019). Blind image super-resolution with spatially variant degradations. *ACM Transactions on Graphics (TOG)*, *38*(6), 1-13.

https://doi.org/10.1145/3355089.3356575

[9] Aravkin, Y. A., Burke, V. J., & Chiuso, A., et al. (2014). Convex vs non-convex estimators for regression and sparse estimation: the mean squared error properties of ARD and GLasso. *Journal of Machine Learning Research*, *15*(1), 217-252.

[10] Wu, Y. C., Hayashi, T., & Tobing, P. L., et al. (2021). Quasi-Periodic WaveNet: An Autoregressive Raw Waveform Generative Model with Pitch-dependent Dilated Convolution Neural Network. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *29*, 1134-1148. https://doi.org/10.1109/TASLP.2021.3061245

[11] Saxena, D. & Cao, J. (2021). Generative Adversarial Networks (GANs): Challenges, Solutions, and Future Directions. *ACM Computing Surveys*, *54*(3), 1-42. https://doi.org/10.1145/3446374

[12] Geng. J., Shao, T., & Zheng, Y., et al. (2018). Warp-guided GANs for single-photo facial animation. *ACM Transactions on Graphics (TOG)*, *37*(6), 1-12. https://doi.org/10.1145/3272127.3275043

[13] Jiang, J., Wang, C., & Liu, X., et al. (2021). Deep Learning-based Face Super-resolution:A Survey. *ACM Computing Surveys (CSUR)*, *55*(1), 1-36. https://doi.org/10.1145/3485132

[14] Papamakarios, G., Nalisnick, E., & Jimenez Rezende, D. et al. (2019). Normalizing Flows for Probabilistic Modeling and Inference. *The Journal of Machine Learning Research (JMLR)*, *22*(1), 2617-2680.

[15] Haque, R., Ho, S. B., Chai, I., & Abdullah, A. (2023). Improved Adam-based feedforward deep neural network model for personalized asthma predictions. *Journal of System an Management Sciences*, *13*(2), 241-257. https://doi.org/10.33168/JSMS.2023.0217

[16] Oord, A., Kalchbrenner, N., & Kavukcuoglu, K. (2016). Pixel Recurrent Neural Networks. *ICML*, *48*, 1747-1756.

[17] Kingma, D. P. & Dhariwal, P. (2018). Glow: Generative Flow with Invertible 1x1 Convolutions. *NIPS*, *1*, 0236-10245.

[18] Madhavi, S. & Hong, S. P. (2022). Anomaly detection using deep neural network quantum encoder. *Journal of Logistics. Informatics and Service Science*, *9*(2), 118-130.

[19] Müller, T., Mcwilliams, B., & Rousselle, F., et al. (2018). Neural Importance Sampling. *ACM Transactions on Graphics*, *38*(5), 1-19. https://doi.org/10.1145/3341156

[20] Zhang, W., Chang, X., & Qian, Y., et al. (2020). Improving End-to-End Single-Channel Multi-Talker Speech Recognition. *IEEE/ACM Transactions on Audio Speech and Language Processing*, *28*, 1385-1394. https://doi.org/10.1109/TASLP.2020.2988423

[21] Badea, L., Moraru, Ad., Ilie, C., Duhnea, C., & Panait, I. (2021). Customer Satisfaction with Banking Services. *Simulatingthe Influence of Customer Satisfaction Determinants Using Artificial Neural Network Algorithms. Economic Computation And Economic Cybernetics Studies And Research*, *55*(4), 101-116, https://doi.org/10.24818/18423264/55.4.21.07

[22] Rezende, D. J., Mohamed, S., & Wierstra, D. (2014). Stochastic Backpropagation and Approximate Inference in Deep Generative Models. *ICML*, *32*, 1278-1286.

[23] Yi, T-y., Hu, P-h., & Liu, Z. (2023). Superre solution Method of ISAR Image Based on Improved CycleGAN. *Signal Processing*, *39*(02), 323-334.

[24] Li, X., Sun, Y., & Yang, Y., et al. (2019). Symmetrical Residual Connections for Single Image Super-Resolution. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, *15*(1), 1-10. https://doi.org/10.1145/3282445

[25] Ho, J., Jain, A., & Abbeel P. (2020). Denoising Diffusion Probabilistic Models. *NIPS*, 6840-6851.

[26] Liu, M., Xue, Z., & Xu, X., et al. (2019). Host-Based Intrusion Detection System with System Calls: Review and Future Trends. *ACM Comput. Surv*, *51*(5), 1-36. https://doi.org/10.1145/3214304

[27] Xia, X. & Li, T. (2019). A fuzzy control model based on BP neural network arithmetic for optimal control of smart city facilities. *Personal and Ubiquitous Computing*, *23*(3-4), 453-463. https://doi.org/10.1007/s00779-019-01209-0

[28] Khan, A. M. & Patel, S. V. (2018). A Simple Modal Logic for Reasoning in Multigranulation Rough Set Model. *ACM Transactions on Computational Logic (TOCL)*, *19*(4), 1-23. https://doi.org/10.1145/3274664

[29] Meng, J., Zhang, J., & Luan, Y. (2015). Gene selection integrated with biological knowledge for plant stress response using neighborhood system and rough set theory. *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, *12*(2), 433-444. https://doi.org/10.1109/TCBB.2014.2361329

[30] Ho, J., Saharia, C., Chan, W., et al. (2021). Cascaded Diffusion Models for High Fidelity Image Generation. *The Journal of Machine Learning Research (JMLR)*, *23*(1), 2249-2281.

[31] Meng, J., Zhang, J., Li, R., & Luan, Y. (2014). Gene selection using rough set based on neighborhood for the analysis of plant stress response. *Applied Soft Computing*, *25*, 51-63. https://doi.org/10.1016/j.asoc.2014.09.013

Contact information:

Tao SONG, Master lecturer
(Corresponding author)
School of Intelligence Technology,
Geely University of China,
Chengdu Sichuan, 610000, P. R. China
No. 123, SEC. 2, Chengjian Avenue,
Eastern New District, Chengdu City, Sichuan Province
E-mail: songtao@guc.edu.cn

Ran WEN, Master associate professor
School of Intelligence Technology,
Geely University of China,
Chengdu Sichuan, 610000, P. R. China
No. 123, SEC. 2, Chengjian Avenue,
Eastern New District, Chengdu City, Sichuan Province
E-mail: wenran@guc.edu.cn

Lei ZHANG, PhD professor
a. College of Business Administration Zhejiang University of Finance and Economics Hangzhou 310018, P. R. China
b. School of Economics and Management Xinjiang University,
Wulumuqi 830046, P. R. China
E-mail: leinuozhang@zufe.edu.cn