

Recent Advances in Image Super-Resolution Reconstruction Based on Machine Learning

Qingru Zhang 

School of Artificial Intelligence, Beijing Normal University, Beijing 100875, China

Corresponding Author Email: 202311081088@mail.bnu.edu.cn

Copyright: ©2024 The author. This article is published by IIETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.410627>

Received: 9 June 2024

Revised: 28 September 2024

Accepted: 17 October 2024

Available online: 31 December 2024

Keywords:

image super-resolution (SR) reconstruction, machine learning, Markov Decision Process (MDP), inter-region reward function, visual coherence

ABSTRACT

Image super-resolution (SR) reconstruction is a crucial research area in computer vision, aiming to restore high-resolution images from low-resolution inputs, thereby enhancing image detail and quality. With the continuous growth of digital image applications, SR technology has been widely utilized in fields such as medical imaging, satellite remote sensing, surveillance video enhancement, and virtual reality. However, despite significant progress in objective image quality, existing SR methods still face challenges such as loss of image details, unnatural textures, and visual inconsistencies. This is especially evident in complex scenes or high-noise environments, where traditional unified models are ineffective in addressing the differences between image regions, resulting in suboptimal reconstruction outcomes. In recent years, deep learning methods, such as Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs), have made remarkable strides in the field of SR. However, most methods still overlook the spatial dependencies between different regions of the image. To address this limitation, this paper proposes a SR reconstruction framework based on the Markov Decision Process (MDP) and Deep Q-Networks (DQN), which dynamically selects SR models using reinforcement learning principles for adaptive optimization across image regions. Furthermore, a new reward function is introduced to resolve the model selection consistency issue across regions, aiming to improve the visual transition between adjacent regions and enhance the overall perceptual quality of the image. Experimental results demonstrate that the proposed framework effectively improves the reconstruction performance of SR images, significantly enhancing visual coherence while maintaining objective quality.

1. INTRODUCTION

Image SR aims to restore high-resolution images from low-resolution images in order to recover the image's details and texture information [1-4]. With the rapid development of digital image processing technology and computer vision applications, image SR reconstruction technology has achieved significant applications in many fields, including medical imaging, satellite remote sensing, surveillance video enhancement, and virtual reality [2, 5-7]. However, despite some progress in traditional SR methods in terms of objective quality, such as improvements in Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity (SSIM) indices, there are still issues like detail loss, unnatural textures, and visual incoherence in the reconstructed images. These problems become more pronounced, especially in complex scenes and high-noise environments [8-10]. Therefore, improving not only the objective quality of SR images but also the perceptual quality for human eyes has become a core issue that needs to be solved in the current field of image SR.

The significance of researching image SR reconstruction lies not only in improving image quality but also in advancing the practical application of related technologies. For example, in medical imaging, SR technology can help doctors more clearly identify the location and shape of lesions; in security

surveillance, SR images can help improve the accuracy of face recognition and license plate recognition; in virtual reality and augmented reality applications, high-quality image reconstruction can enhance immersion and interactive experiences [11-15]. In recent years, with the rise of deep learning, SR methods based on CNNs and GANs have made breakthrough progress, particularly in handling more complex image textures and finer details [3, 16]. However, existing methods often overlook the spatial dependencies between image regions, making it difficult to balance local and global features, thus affecting the final visual coherence and overall effect of the image.

Current SR technologies mainly rely on unified models to process the entire image, but different regions of the image often have different textures, structures, and noise characteristics. Therefore, a single model is often unable to meet the needs of all regions [17-21]. In addition, traditional training methods based on loss functions (such as Mean Squared Error, MSE) can improve the objective quality of images but often neglect visual consistency, leading to unnatural transitions and distorted details. Regarding these issues, existing methods have failed to effectively combine the interrelationships between image regions and the consistency of model selection, resulting in significant subjective visual differences in image reconstruction. Therefore, exploring a

more refined SR reconstruction framework and incorporating compatibility considerations between regions in model selection has become an important direction in current research.

This paper mainly includes two aspects: first, a new image SR reconstruction framework is proposed based on MDP and machine learning. This framework utilizes reinforcement learning to dynamically select SR models and adaptively optimize the reconstruction process in different image regions. Second, to address the local inconsistency issue in existing SR model selection, this paper proposes an improved method for the inter-region reward function, aiming to improve the consistency of model selection between adjacent image regions and ensure visual coherence in image reconstruction. Through these two innovations, this paper can significantly improve the subjective perceptual quality of image SR reconstruction while enhancing the objective performance, thus providing theoretical support and technical guarantees for the promotion and application of SR technology in practical applications.

2. IMAGE SR RECONSTRUCTION BASED ON MDP AND MACHINE LEARNING

Medical images typically contain multiple different regions, some of which are crucial for diagnosis and require extremely high clarity and accuracy. Other regions may be relatively simple, and lower resolution does not affect diagnostic performance. Remote sensing images, which usually come from satellites or drones, have extremely complex surface features, including various textures such as urban areas, forests, lakes, and mountains. The resolution requirements for these regions vary. For example, urban areas may need precise reconstruction of building and road details, while natural landscapes such as forests and lakes may have lower resolution requirements. For these types of images, traditional SR methods often fail to flexibly target the needs of different regions, which may lead to wasted computational resources or imbalanced reconstruction quality. To address this, this paper proposes a new framework for image SR reconstruction with dynamic region selection. The core innovation of this framework lies in the use of reinforcement learning to

dynamically select the most suitable SR model for each texture region, rather than simply applying a fixed SR algorithm to the entire image.

This method is effectively modeled through a reinforcement learning framework, where MDP provides a clear theoretical structure. The texture characteristics of different organs or tissues in medical images often vary significantly, and the texture and scale variations of different ground objects in remote sensing images are also very complex. In this framework, the input low-resolution image is divided into several sub-blocks through a segmentation algorithm, each containing different texture features. The goal is to select the most suitable model from a predefined pool of SR models based on these features and apply it to each image sub-block. This approach flexibly addresses the regional heterogeneity in images such as those from medical imaging and remote sensing, improving the accuracy and application of SR reconstruction. By modeling with MDP, the image SR task can be treated as a sequential decision problem, where each state represents the current state of the image, each action is the selection of the SR model to be applied to a specific image sub-block, and the reward function provides feedback based on the MSE between the region's SR reconstruction and the Groundtruth image.

To allow the reinforcement learning model to dynamically select the most suitable SR model between different image sub-blocks, the action space of MDP is defined as a set of J-dimensional one-hot vectors. In MDP, each image sub-block is a state, and the action selection is determined based on the specific needs and features of each state. The ultimate goal is to optimize the policy so that the overall SR reconstruction effect of the image is optimal. Whenever a specific SR model is chosen to process a particular image sub-block, this action has an immediate effect on a portion of the image, and this effect is fed back to the agent via the reward function. During training, the agent aims to maximize long-term rewards based on the current state and action selection. This requires considering not only the effect of processing the current image sub-block but also taking into account the final reconstruction quality of the entire image. Through this dynamic selection mechanism, MDP can effectively select the best SR model for each image sub-block, thereby improving the overall effect of the SR reconstruction.

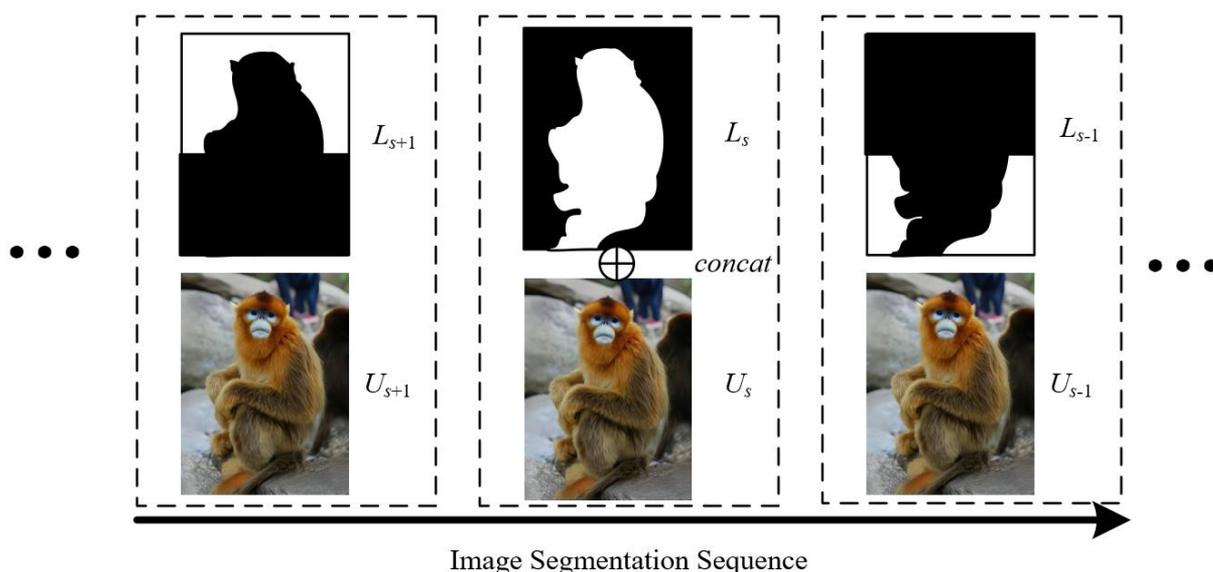


Figure 1. Example of image state transition

In the dynamic selection modeling of the SR model in this paper, the state of MDP determines how the model dynamically selects SR models between different image sub-blocks. To ensure that each state maintains physical significance during the state transition process, the state needs to contain both global and local information of the image. Specifically, in this implementation, the state is defined as consisting of two parts: on one hand, the entire image processed by the SR model, i.e., the image processed in the current iteration; on the other hand, the segmentation label corresponding to the current image sub-block, which defines the image region currently being processed. During the training process, the state is formed by concatenating these two elements, ensuring that the state contains both the global information of the entire image and the local information of the current image sub-block. With this state definition, the state transition in MDP is effectively guaranteed. In the MDP framework, every time an action is performed, the image and state will change. For example, when the agent selects a certain SR model to process the current image sub-block, the image state will change, and the entire image will transition from the current state to the next state, i.e., the current image sub-block is processed and updated to the next image sub-block. This process ensures that the state transitions are meaningful, not independent, and unfold gradually as image sub-blocks are processed. After each iteration, the new state will include the updated image information U^s and the segmentation label L_s of the next image sub-block, thereby maintaining the continuous connection between global information and local information. This state transition method ensures that the reinforcement learning model can maintain a dynamic balance between global and local information throughout the training process, enabling the selected SR model to provide the best reconstruction effect for each sub-block and optimize the final SR image quality through continuous state transitions. Figure 1 provides an example of image state transition. The current state is defined as:

$$T_s = U^s \oplus L_s \quad (1)$$

Assuming the negation of L_s is represented as \bar{L}_s , the SR

image after processing the s -th image sub-block using the action x_s is represented by $d(U^s)$, and pixel-level multiplication is denoted by $*$; the state transition from U^s to U^{s+1} can be represented as:

$$U^{s+1} = \bar{L}_s * U^s + L_s * d(U^s | x_s) \quad (2)$$

Based on the above two formulas, the state transition definition is as follows:

$$T_{s+1} = [\bar{L}_s * U^s + L_s * d(U^s | x_s)] \oplus L_{s+1} \quad (3)$$

In the dynamic decision-making framework based on reinforcement learning, the reward function evaluates the effect of each action to provide feedback on the quality of the SR model selected by the agent, thus optimizing the policy. Specifically, for the image SR task, although the agent does not know which SR model is optimal for each image sub-block at each iteration, by designing a suitable reward function, the agent can gradually learn to select the most suitable model. The reward function proposed in this paper mainly evaluates the reconstruction quality of the current image region and the overall image, with the goal of pushing the agent to select the best SR model by measuring the reconstruction effect of each image sub-block. To achieve this goal, the regional reward function $E_Q(U^s, x_s)$ is designed to measure the difference between the image region processed by the current SR model and the original high-resolution image, with the commonly used metric being the MSE. Mathematically, assuming the standard reinforcement learning reward value is denoted by λ , the threshold is denoted by ς , the Groundtruth high-resolution image is denoted by U^{GE} , and the image width and height of the target high-resolution image are denoted by Q and G , respectively, E_Q is defined as:

$$E_Q(U^s, x_s) = \begin{cases} \lambda & \frac{1}{QG} \|U^{GE} - d(U^{s-1} | x_s)\|_2 < \varsigma \\ -\lambda & otherwise \end{cases} \quad (4)$$

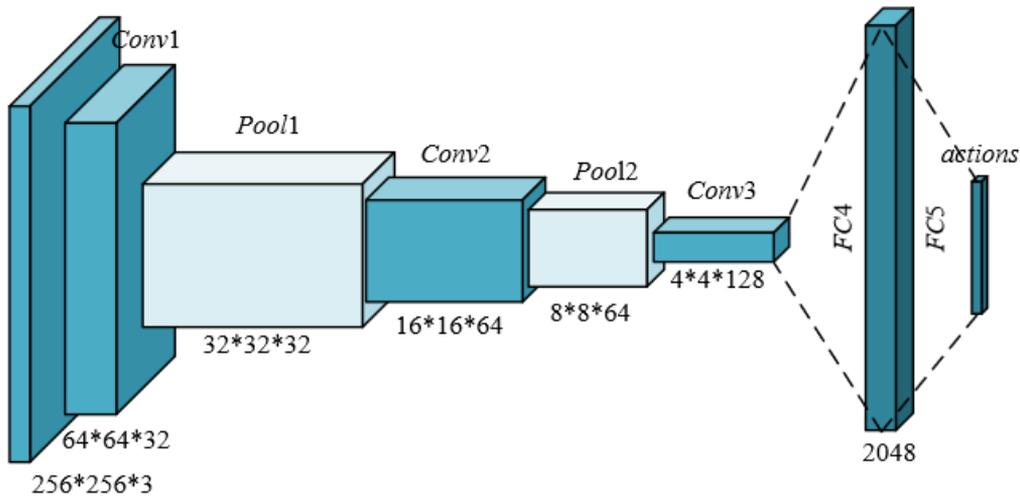


Figure 2. DQN framework

In the SR task, the core goal of the dynamic policy selection framework based on reinforcement learning is to implement dynamic selection of SR models for different image regions through the DQN. The policy in the framework, i.e., the agent, will select the appropriate action, i.e., the SR model, to process the current image sub-block based on the currently observed state. Specifically, the DQN network continuously receives feedback from the environment, i.e., the reward function, and in each iteration, selects the optimal SR model based on the current state, thereby optimizing the overall SR reconstruction quality of the image. In each round of training, the DQN network will, during the processing of each image sub-block, execute an action to select a pre-trained SR model from the model pool, and update the network weights based on the image reconstruction effect produced by the currently selected model. Figure 2 shows the structure of the DQN used.

In the training phase of the DQN, the agent selects the appropriate action x_s based on observing the current state $T_s = U^s \oplus L_s$, and performs dynamic selection of the SR model. According to the ϵ -greedy strategy, the agent randomly selects an action with probability ϵ , which ensures the exploration of more action space, preventing the model from converging to a local optimum during the early stages of training when the network parameters are not fully refined; while with probability $1-\epsilon$, the agent selects the optimal action predicted by the DQN network, which enhances the model's exploitation ability. After selecting the action, the agent calculates the reward of the current action based on the regional reward function $E(D, X)$ designed in this paper, and then updates the Q-value function in the Q-network based on this feedback. This process updates the parameters based on the Bellman equation, and through Q-value optimization, the DQN gradually learns how to select the most suitable SR model for each image region.

The Q-learning network structure consists of three convolutional layers and two fully connected layers. Each convolutional layer uses the ReLU activation function, which helps introduce non-linear features and allows the network to effectively learn complex image features. In the convolution operation, the kernel sizes are 8×8 , 4×4 , and 3×3 , aimed at capturing local information at different scales. By using max-pooling layers for dimensionality reduction, the network is able to compress the input feature map in the first two layers, retaining key information while reducing computation. After each action selection and reward acquisition, the DQN network calculates the current Q-value based on the current state T_s and selected action x_s , and compares it with the target Q-value obtained from the environment feedback. Let the discount factor be represented by η , the probability distribution of the state T_s be represented by γ , and the probability distribution of the action be represented by $\vartheta(\cdot)$. The target Q-value is estimated using the next state T_{s+1} and the maximum Q-value of the selected action, i.e.:

$$W^*(t, x) = R_{t-r} \left[e_s + \eta \text{MAX}_x W^*(t', x') | t, x \right] \quad (5)$$

$$M_s(\varphi_s) = R_{t, x \sim \vartheta(\cdot)} \left[\left(b_s - W(T, x; \varphi_s) \right)^2 \right] \quad (6)$$

$$b_s = \begin{cases} e_s & \text{FO MIN ALS}_{s+1} \\ e_s + \epsilon \text{MAX}_x W(t_{s+1}, x'; \varphi) & \text{otherwise} \end{cases} \quad (7)$$

The loss function of the Q-network is:

$$\nabla_{\varphi_s} M_s(\varphi_s) = R_{t, x \sim \vartheta(\cdot); t'_s \sim \gamma} \left[\begin{array}{c} e_s + \\ \eta \text{MAX}_{x_s} W(t'_s, x'_s; \varphi_{s-1}) \\ -W(t_s, x_s; \varphi_s) \\ \nabla_{\varphi_s} W(t_s, x_s; \varphi_s) \end{array} \right] \quad (8)$$

3. REWARD FUNCTION IMPROVEMENT FOR VISUAL TRANSITION BETWEEN REGIONS

Due to the similarities in texture and structure between image regions, selecting inconsistent or incompatible SR models may result in noticeable stitching marks or visual incoherence at the image boundaries, affecting the overall subjective quality of the image. Therefore, traditional reward functions based on MSE ignore the spatial dependencies between regions, leading to a lack of natural and smooth transitions in the SR image. Although objective metrics may be high, the best visual performance is not achieved. To overcome this issue, the reward function in this paper is redesigned to not only consider the MSE between the current image region and the ground truth, as in the traditional MSE part, but also to fully consider the collaborative nature between different SR models, especially the consistency of model selection between adjacent regions.

In specific tasks like medical image and remote sensing image SR reconstruction, the improvement of image quality relies not only on the restoration of local details but also on maintaining the overall consistency and naturalness of the image. The reward function improvement in this paper mainly targets the consistency of SR model selection between different regions, especially when dealing with artifacts and unnatural texture transitions at the boundary regions. In medical images, especially at the boundaries of organs or tissues, if different SR models are used for adjacent regions, it may lead to obvious discontinuities or gaps in the image, which could impact the accurate judgment of the lesion area by doctors. In remote sensing images, frequently switching SR models may cause inconsistent detail restoration in different texture regions of the ground features, even resulting in false boundaries that affect the accuracy of feature recognition. Therefore, the improved reward function not only selects the most suitable SR model for each region but also introduces consistency constraints, preventing the use of too many different models within the same image, thereby ensuring the overall perceptual quality of the image. Especially in regions with significant texture changes, reducing model switching can effectively prevent over-rendering of image details or unnatural texture variations, enhancing the final reconstructed image's visual effect and making it more suitable for medical diagnosis and remote sensing analysis.

For the above reasons, this paper proposes the inter-region reward function $E_Y(U^s, x_s)$. Assume that the number of SR models in the model pool is represented by J . The number of image sub-blocks contained in the segmented input image is represented by V . The indicator matrix of size J^*V is represented by O , the indicator function by $\sigma(\cdot)$, and the logical 'or' by the symbol N . The weight coefficient is represented by α , and its specific expression is as follows:

$$E_Y(U^s, x_s) = E_Z(U^s, x_s) + \alpha E_V(U^s, x_s) \quad (9)$$

$$E_Z(U^s, x_s) = -\sum_{j=1}^J \sum_{u \in V(k)} \sigma(O(j, u), O(j, k)) \quad (10)$$

$$E_V(U^s, x_s) = -\sum_{j=1}^J \bigvee_{u=1}^V O(j, u) \quad (11)$$

Figure 3 compares the texture details in SR images reconstructed with the reward function before and after improvement.



Figure 3. Comparison of texture detail in SR images reconstructed with the reward function before and after improvement

4. EXPERIMENTAL RESULTS AND ANALYSIS

Table 1 shows the comparison of reconstruction quality before and after the improvement, based on different SR scales ($\times 2$, $\times 3$, $\times 4$) for the four datasets: Urban100, T91, DIV2K, and Flickr2K. The evaluation metrics include PSNR and IFC (Structural Similarity Index). From the data before improvement, the PSNR values for the four datasets at the $\times 2$ scale are 36.25, 32.15, 31.25, and 32.24, and the corresponding IFC values are 8.569, 8.235, 7.785, and 8.785. At the $\times 3$ scale, the PSNR values decrease to 32.15, 31.25, 27.88, and 26.35, with the IFC values dropping to 5.265, 4.562, 4.125, and 5.231. The results at the $\times 4$ scale further degrade. These results indicate that without the improved reward function, the image quality sharply declines as the SR scale increases, particularly for datasets 3 and 4, where both PSNR and IFC values show significant degradation. In contrast, the algorithm with the improved reward function demonstrates a noticeable quality improvement, especially at the $\times 3$ and $\times 4$ scales. For example, at the $\times 3$ scale, the PSNR for dataset 1 increases from 32.15 to 33.24, and the IFC improves from 5.265 to 5.214. For dataset

2, the PSNR increases from 31.25 to 31.25, and the IFC improves from 4.562 to 4.568. Additionally, although the improvement for datasets 3 and 4 is relatively weaker, the PSNR and IFC values still show slight improvements, especially at the $\times 4$ scale.

The experimental results show that the proposed improved reward function significantly enhances the image SR reconstruction quality, especially for large-scale upscaling ($\times 3$ and $\times 4$), where the effect is more pronounced. Specifically, the improved inter-region reward function effectively alleviates the local inconsistency issue in model selection, enhancing the visual coherence of the image reconstruction by improving the consistency of model selection between adjacent regions. PSNR and IFC values for datasets 1 and 2 show stable and significant improvements at various scales, indicating that the improved algorithm can enhance overall image quality while preserving detail. For datasets 3 and 4, although the improvement is relatively modest, the increase in PSNR and IFC values still demonstrates the adaptability and effectiveness of the improvement method across different types of images.

Based on the experimental data from Tables 2, 3, and 4, significant differences in the performance (such as PSNR and SSIM) and model complexity (such as parameter count and Multi-Adds) of different algorithms are observed for the SR reconstruction task. Firstly, regarding performance metrics, EDSR and RCAN show consistent results across scales $\times 2$, $\times 3$, and $\times 4$, achieving high PSNR values (around 31.26) and SSIM values (ranging from 0.9125 to 0.9245). This indicates that these models are effective in recovering image details and maintaining visual quality. In contrast, while SRCNN and VDSR perform reasonably well at scale $\times 2$ (PSNRs of 28.65 and 28.97, SSIMs of 0.8895 and 0.9125), their performance declines significantly as the scaling factor increases. Especially at scales $\times 3$ and $\times 4$, the PSNR values drop markedly, indicating that these algorithms struggle with higher magnification compared to other deep learning methods. Secondly, in terms of model complexity, the proposed method is relatively optimized in terms of parameter count and Multi-Adds compared to other common SR algorithms (like SRCNN, VDSR, and FSRCNN). For instance, at scale $\times 2$, SRCNN and VDSR have parameter counts of 8K and 12K, while the proposed method has 558K parameters, which is a moderate increase compared to traditional algorithms. However, for scales $\times 3$ and $\times 4$, the proposed method's parameter count is reasonably controlled, being notably lower than EDSR (536K and 546K at scales $\times 3$ and $\times 4$, respectively). Additionally, the proposed method shows good optimization in terms of Multi-Adds. For scales $\times 3$ and $\times 4$, the proposed method has Multi-Adds of 53.2G and 32.6G, respectively, which is lower than traditional deep networks like FSRCNN (72.3G) and EDSR (41.2G). This demonstrates the computational efficiency advantage of the proposed model.

Table 1. Comparison of reconstruction quality with the proposed algorithm before and after reward function improvement

Dataset	Scale	Dataset 1 PSNR/IFC	Dataset 2 PSNR/IFC	Dataset 3 PSNR/IFC	Dataset 4 PSNR/IFC
Before Improvement	$\times 2$	36.25/8.569	32.15/8.235	31.25/7.785	32.24/8.785
	$\times 3$	32.15/5.265	31.25/4.562	27.88/4.125	26.35/5.231
	$\times 4$	31.24/3.698	27.89/3.125	26.58/2.658	24.15/3.562
After Improvement	$\times 2$	36.25/8.569	32.65/8.124	31.25/7.785	32.15/8.795
	$\times 3$	33.24/5.214	31.25/4.568	27.15/4.125	26.25/5.231
	$\times 4$	31.26/3.654	27.26/3.125	26.58/2.658	24.56/3.568

Table 2. Performance and model complexity comparison of different algorithms at scale $\times 2$

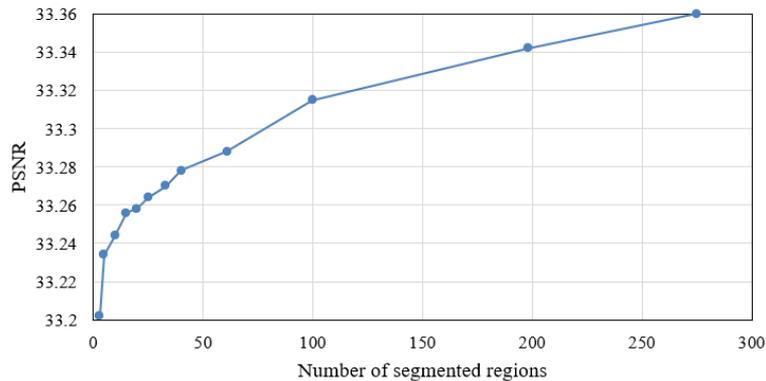
Method ($\times 2$)	<i>SRCNN</i>	<i>VDSR</i>	<i>FSRCNN</i>	<i>EDSR</i>	<i>RCAN</i>	<i>ESRGAN</i>	<i>VDSR</i>	Proposed Method
Parameters	8	12	689	526	865	735	578	558
<i>Multi-Adds</i>	51.6	6	157.8	94	189.26	169	156.9	132.8
<i>PSNR</i>	28.65	28.97	31.26	31.26	31.26	31.26	31.25	31.25
<i>SSIM</i>	0.8895	0.9125	0.9125	0.9125	0.9245	0.9236	0.9236	0.9236

Table 3. Performance and model complexity comparison of different algorithms at scale $\times 3$

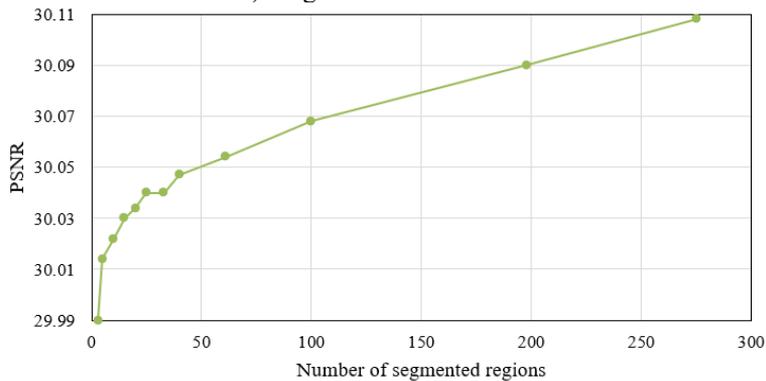
Method ($\times 3$)	<i>SRCNN</i>	<i>VDSR</i>	<i>FSRCNN</i>	<i>EDSR</i>	<i>RCAN</i>	<i>ESRGAN</i>	<i>VDSR</i>	Proposed Method
Parameters	8	12	712	536	878	745	588	569
<i>Multi-Adds</i>	51.2	5.2	72.3	41.2	86.2	76.2	74.2	53.2
<i>PSNR</i>	25.69	25.36	27.56	27.89	27.56	27.56	26.36	27.15
<i>SSIM</i>	0.7895	0.8123	0.8452	0.8456	0.8562	0.8456	0.8569	0.8562

Table 4. Performance and model complexity comparison of different algorithms at scale $\times 4$

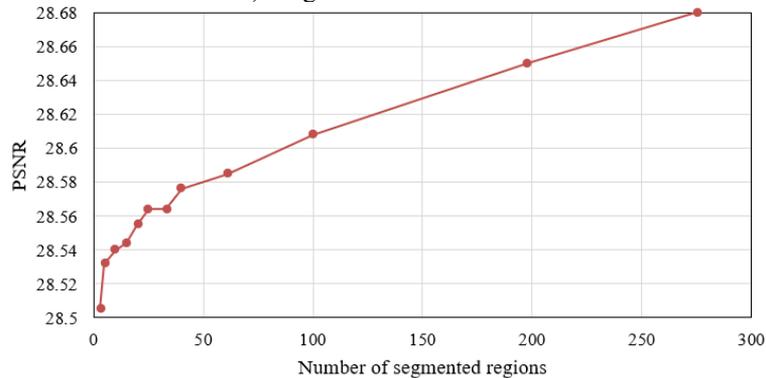
Method ($\times 4$)	<i>SRCNN</i>	<i>VDSR</i>	<i>FSRCNN</i>	<i>EDSR</i>	<i>RCAN</i>	<i>ESRGAN</i>	<i>VDSR</i>	Proposed Method
Parameters [K]	8	12	725	546	889	756	612	578
<i>Multi-Adds</i> [G]	51.2	4.5	41.2	22.6	48.5	43.2	42.5	32.6
<i>PSNR</i>	23.25	23.55	25.21	25.69	25.68	25.58	25.69	25.15
<i>SSIM</i>	0.7152	0.7125	0.7789	0.7789	0.7895	0.7894	0.7956	0.8123



1) Magnification factor of 2



2) Magnification factor of 3



3) Magnification factor of 4

Figure 4. Curves showing the change in PSNR with varying numbers of sub-image regions for different magnification factors

The experimental results suggest that the proposed method outperforms traditional SRCNN and VDSR methods, particularly at higher magnification ($\times 3$ and $\times 4$). It strikes a better balance between performance and complexity. Specifically, at scales $\times 3$ and $\times 4$, the proposed method achieves PSNR and SSIM values close to those of more complex models like RCAN and EDSR, but with significantly lower computational complexity. This indicates that the proposed SR framework, based on MDP and reinforcement learning, provides a practical solution for dynamically selecting models and optimizing the reconstruction process, particularly for adaptive handling of different image regions. It ensures high image quality while reducing computational load.

According to the results provided in Figure 4, as the magnification factor increases ($\times 2$, $\times 3$, and $\times 4$), the PSNR value gradually improves with an increase in the number of sub-image regions. For a magnification factor of 2, as the number of sub-image regions increases from 3 to 100, the PSNR value increases from 33.2025 to 33.362, indicating a noticeable improvement in image reconstruction quality with finer region segmentation. Similarly, for a magnification factor of 3, the PSNR value increases from 29.995 with 3 regions to 30.108 with 198 regions, further reflecting the impact of region division on reconstruction performance. A similar trend is observed at a magnification factor of 4, where the PSNR value starts at 28.5055 and gradually rises to 28.68. Although the PSNR values differ across magnification factors, in each case, as the number of regions increases, the PSNR value shows a gradual optimization trend. This trend aligns with the SR reconstruction framework proposed in this paper, which is based on MDP and machine learning. Specifically, the introduction of reinforcement learning enables the model to dynamically select appropriate reconstruction methods based on different regions of the image, significantly improving SR reconstruction quality. As the number of sub-image regions increases, the model can optimize smaller regions, improving the accuracy of image reconstruction, especially in complex regions with fine details.

From the experimental results, it can be seen that the proposed dynamic model selection method based on reinforcement learning effectively improves the image reconstruction quality at different magnification factors by increasing the number of sub-image regions. Through the improvement of the reward function between regions, the proposed method can make more precise model selections within local regions, reducing the local inconsistency issues commonly found in traditional SR models. As the number of regions increases, the PSNR value gradually improves, indicating that the consistency of model selection between adjacent image regions has been optimized, and the visual coherence of the image reconstruction has been ensured.

5. CONCLUSION

This paper proposes an innovative image SR reconstruction framework that combines MDP and machine learning methods to improve the quality and computational efficiency of image SR reconstruction. Specifically, by introducing the concept of reinforcement learning, this paper dynamically selects SR models for different regions of the image and adaptively optimizes the reconstruction process based on the characteristics of each region. The core innovation of this

method lies in dynamically adjusting the reconstruction strategy, enabling the model to select the most suitable SR model for each region based on the changing image content, thereby improving detail recovery and global image quality. The research presented in this paper has significant academic and practical value in the field of image SR. By combining MDP and reinforcement learning, a new SR reconstruction framework is proposed, which achieves dynamic model selection and adaptive optimization for different regions of the image. This framework breaks the limitations of traditional SR methods, making the model more flexible and efficient when processing different regions of the image. Additionally, the introduction of the reward function between regions successfully solves the common issue of local inconsistency in SR models, improving the visual coherence and naturalness of the image. These innovations not only advance theoretical research but also provide more effective solutions for image reconstruction problems in practical applications.

Although the proposed method demonstrates good performance in several aspects, there are still certain limitations. Firstly, although image quality is improved by region segmentation, further enhancing the model's adaptability and the precision of its selection strategy remains a challenge for extremely complex or highly detailed images. Secondly, compared to some lightweight models, the current method still has a certain gap in computational efficiency, especially when processing large-scale high-resolution images, which leads to higher computational costs. Further optimization of the algorithm's speed and efficiency is needed. Moreover, the model's adaptability to different image content may need to be tested on more real-world datasets to verify its generalization and robustness. Future research could focus on the following directions: first, further optimizing the reward design in reinforcement learning and the MDP to make the model more intelligent in handling SR tasks in complex scenarios. Second, the proposed method could be combined with lightweight network architectures to reduce computational complexity and improve real-time processing capabilities. Another direction is to explore cross-scale and cross-domain SR reconstruction methods, enabling the model to not only adapt to images with different resolutions but also handle reconstruction tasks for different types of images (such as dynamic images or video frames). Finally, expanding the application of the reward function between regions to explore how to achieve consistency across larger image regions would help better handle the detail transition issues in complex scenes.

REFERENCES

- [1] Hu, Y., Lam, K.M., Qiu, G., Shen, T. (2010). From local pixel structure to global image super-resolution: A new face hallucination framework. *IEEE Transactions on Image Processing*, 20(2): 433-445. <https://doi.org/10.1109/TIP.2010.2063437>
- [2] Jiang, L., Ye, S., Zhao, L., Ma, X., Yang, X. (2019). Medical image super-resolution for remote medical diagnosis in smart city: A case study based on the new healthcare reform of China. *Sustainable Cities and Society*, 48: 101497. <https://doi.org/10.1016/j.scs.2019.101497>
- [3] Xu, S.H., Qi, M.M., Wang, X.M., Zhao, H.L., Hu, Z.Y., Sun, H.Y. (2022). A positive-unlabeled generative

- adversarial network for super-resolution image reconstruction using a Charbonnier loss. *Traitement du Signal*, 39(3): 1061-1069. <https://doi.org/10.18280/ts.390333>
- [4] Temiz, H. (2023). Enhancing the resolution of historical Ottoman texts using deep learning-based super-resolution techniques. *Traitement du Signal*, 40(3): 1075-1082. <https://doi.org/10.18280/ts.400323>
- [5] Huang, H., He, R., Sun, Z., Tan, T. (2019). Wavelet domain generative adversarial network for multi-scale face hallucination. *International Journal of Computer Vision*, 127(6): 763-784. <https://doi.org/10.1007/s11263-019-01154-8>
- [6] Rajput, S.S. (2022). Mixed gaussian-impulse noise robust face hallucination via noise suppressed low-and-high resolution space-based neighbor representation. *Multimedia Tools and Applications*, 81(11): 15997-16019. <https://doi.org/10.1007/s11042-022-12154-1>
- [7] Rajput, S.S., Singh, A., Arya, K.V., Jiang, J. (2018). Noise robust face hallucination algorithm using local content prior based error shrunk nearest neighbors representation. *Signal Processing*, 147: 233-246. <https://doi.org/10.1016/j.sigpro.2018.01.030>
- [8] Lee, A., Tsekouras, K., Calderon, C., Bustamante, C., Pressé, S. (2017). Unraveling the thousand word picture: An introduction to super-resolution data analysis. *Chemical Reviews*, 117(11): 7276-7330. <https://doi.org/10.1021/acs.chemrev.6b00729>
- [9] Zareapoor, M., Jain, D.K., Yang, J. (2018). Local spatial information for image super-resolution. *Cognitive Systems Research*, 52: 49-57. <https://doi.org/10.1016/j.cogsys.2018.06.007>
- [10] Mckean, L.N., Newman, E.F., Adair, P. (2013). Feeling like me again: A grounded theory of the role of breast reconstruction surgery in self-image. *European Journal of Cancer Care*, 22(4): 493-502. <https://doi.org/10.1111/ecc.12055>
- [11] Hao, Q., Zheng, W., Wang, C., Xiao, Y., Zhang, L. (2024). MLRN: A multi-view local reconstruction network for single image restoration. *Information Processing & Management*, 61(3): 103700. <https://doi.org/10.1016/j.ipm.2024.103700>
- [12] Li, Q., Wu, F., Chen, G. (2019). An efficient, fair, and robust image pricing mechanism for crowdsourced 3D reconstruction. *IEEE Transactions on Services Computing*, 15(1): 498-512. <https://doi.org/10.1109/TSC.2019.2953906>
- [13] Teo, I., Fronczyk, K.M., Guindani, M., Vannucci, M., Ulfers, S.S., Hanasono, M.M., Fingeret, M.C. (2016). Salient body image concerns of patients with cancer undergoing head and neck reconstruction. *Head & Neck*, 38(7): 1035-1042. <https://doi.org/10.1002/hed.24415>
- [14] Davidi, R., Herman, G.T., Censor, Y. (2009). Perturbation-resilient block-iterative projection methods with application to image reconstruction from projections. *International Transactions in Operational Research*, 16(4): 505-524. <https://doi.org/10.1111/j.1475-3995.2009.00695.x>
- [15] Park, S.W., Yoon, R.G., Lee, H., Lee, H.J., Choi, Y.D., Lee, D.H. (2020). Impacts of thresholds of gray value for cone-beam computed tomography 3D reconstruction on the accuracy of image matching with optical scan. *International Journal of Environmental Research and Public Health*, 17(17): 6375. <https://doi.org/10.3390/ijerph17176375>
- [16] Hopkins, R.O., Abildskov, T.J., Bigler, E.D., Weaver, L.K. (1997). Three dimensional image reconstruction of neuroanatomical structures: Methods for isolation of the cortex, ventricular system, hippocampus, and fornix. *Neuropsychology Review*, 7: 87-104. <https://doi.org/10.1023/B:NERV.0000005946.46506.a9>
- [17] Chang, C.H., Nemrodov, D., Drobotenko, N., Sorkhou, M., Nestor, A., Lee, A.C. (2021). Image reconstruction reveals the impact of aging on face perception. *Journal of Experimental Psychology: Human Perception and Performance*, 47(7): 977-991. <https://doi.org/10.1037/xhp0000920>
- [18] Jang, Y., Seong, M., Sok, S. (2023). Influence of body image on quality of life in breast cancer patients undergoing breast reconstruction: Mediating of self-esteem. *Journal of Clinical Nursing*, 32(17-18): 6366-6373. <https://doi.org/10.1111/jocn.16621>
- [19] Ahn, J., Suh, E.E. (2021). The lived experience of body alteration and body image with regard to immediate breast reconstruction among women with breast cancer. *Journal of Korean Academy of Nursing*, 51(2): 245-259. <https://doi.org/10.4040/jkan.21028>
- [20] Harcourt, D., Russell, C., Hughes, J., White, P., Nduka, C., Smith, R. (2011). Patient satisfaction in relation to nipple reconstruction: The importance of information provision. *Journal of Plastic, Reconstructive & Aesthetic Surgery*, 64(4): 494-499. <https://doi.org/10.1016/j.bjps.2010.06.008>
- [21] Ceballos, L.M., RojasDeFrancisco, L., Osorio, J.C.M. (2020). The role of a fashion spotlight event in a process of city image reconstruction. *Journal of Destination Marketing & Management*, 17: 100464. <https://doi.org/10.1016/j.jdmm.2020.100464>