Research on the Improvement of Image Super Resolution Reconstruction Algorithm Based on AWSRN Model

Bin Dong
School of Computer Science and Engineering
Xi'an Technological University
Xi'an, 710021, Shaanxi, China
E-mail: 2670538455@qq.com

Zhiyi Hu
Engineering Design Institute
Army Research Laboratory
Beijing, 100042, China
E-mail: 763757335 @qq.com

Jun Yu
School of Computer Science and Engineering
Xi'an Technological University
Xi'an, 710021, Shaanxi, China
E-mail: yujun@xatu.edu.cn

Feng Xiong
School of Computer Science and Engineering
Xi'an Technological University
Xi'an, 710021, Shaanxi, China
E-mail: 1023465977@qq.com

***Abstract*—In domains such as medical diagnostics, surveillance technology, and geospatial imaging, the escalating need for ultra-high-definition imagery has exposed the limitations of conventional super-resolution methods. These legacy algorithms often fail to deliver the precision and clarity demanded by modern applications. Therefore, this article proposes an optimization algorithm based on the AWSRN network model, aiming to achieve efficient image super-resolution reconstruction, reduce computational costs, and enhance image realism. Firstly, optimize the internal structure of the network and enhance its feature extraction and fusion capabilities; Secondly, to enhance feature extraction precision, a novel module integrating depth-separable convolution with an attention-based mechanism is proposed. Additionally, a hybrid loss function- merging perceptual quality metrics with adversarial training objectives-is employed to rigorously evaluate the disparity between generated and ground-truth images. The MPTS training strategy further optimizes convergence efficiency. Empirical evaluations demonstrate that the enhanced AWSRN model achieves substantial improvements over its baseline counterpart across multiple upscaling factors, particularly at 4× magnification. Specifically, on the Urban100 benchmark, the proposed method elevates PSNR by 1.06 points and SSIM by 0.0239, while maintaining computational efficiency. These advancements offer valuable insights for high-fidelity image upscaling methodologies.**

Keywords-Deep Learning; Image Super-Resolution Reconstruction; AWSRN Network Architecture; Algorithm Optimization

I. INTRODUCTION

In recent years, image super-resolution enhancement has emerged as a prominent focus in visual data optimization. Its primary objective is to reconstruct high-definition visuals from their degraded low-quality counterparts, addressing the growing need for enhanced image fidelity. This approach demonstrates significant applicability across diverse domains, including medical diagnostics, surveillance systems, and satellite imagery analysis. However, super-resolution reconstruction technology also has its shortcomings. Firstly, the super-resolution reconstruction algorithm is computationally complex and requires strict hardware computing resources, which limits its application in scenarios with high real-time requirements or hardware limitations. Secondly, in terms of texture detail restoration, reconstructed images are prone to issues such as artifacts and blurriness, which may result in discrepancies with high-resolution real images. Thirdly, existing models have poor generalization ability, and models trained on specific datasets have poor reconstruction performance when faced with complex and diverse real-world images.

The field of computer vision has witnessed remarkable advancements in enhancing image

DOI: 10.2478/ijanmc-2025-0015

resolution through deep learning techniques in the past decade. Initial approaches relying on interpolation-based algorithms and conventional modeling techniques demonstrated constrained capabilities, until neural network methodologies revolutionized this domain. A pivotal development occurred when researchers led by Chao Dong introduced SRCNN (Super-Resolution Convolutional Neural Network), marking the inaugural successful implementation of CNN architectures for resolution enhancement tasks. This framework employed direct training to establish nonlinear transformations between degraded and high-quality image spaces, yielding substantially superior results compared to classical approaches [1].

Subsequent architectural innovations continued to push performance boundaries. The VDSR (Very Deep Super-Resolution) architecture, developed in 2016, enhanced output quality through increased network depth, demonstrating measurable improvements in quantitative metrics including peak signal-to-noise ratio, thereby producing outputs with greater fidelity to reference high-definition images [2]. That same year saw the introduction of DRCN (Deeply-Recursive Convolutional Network), which implemented parameter-sharing through recursive connections, achieving comparable accuracy with reduced computational overhead and more efficient resource utilization [3].

Building upon these foundations, subsequent architectural refinements yielded continuous improvements. The 2016-introduced VDSR architecture enhanced reconstruction fidelity through network depth expansion, demonstrating superior performance on quantitative assessment measures including signal-to-noise ratio metrics, thereby producing outputs with enhanced objective quality relative to reference high-definition images [4]. That same year saw the development of DRCN, which employed recursive connectivity patterns to achieve parameter efficiency without compromising reconstruction precision, thereby optimizing computational resource utilization [4].

A significant advancement emerged in 2019 with Chaofeng Wang's team introducing AWSRN, an adaptive learning framework for resolution enhancement. This lightweight architecture incorporated dynamic feature weighting mechanisms that automatically adjusted fusion parameters based on regional importance within the input image, substantially enhancing both output realism and processing efficiency [5]. However, the model exhibits limitations when processing complex scenes or non-standard textures. Particularly for artistic imagery with distinctive color distributions and textural patterns - characteristics often divergent from conventional training datasets - the system frequently generates artifacts including distorted textures and inaccurate color reproduction, ultimately failing to preserve the original stylistic integrity in the enhanced outputs [6].To address these challenges, our study presents architectural refinements to both the adaptive weighting residual components (AWRU) and region-specific feature integration modules (LRFU). These modifications strengthen the network's capacity for hierarchical feature processing within localized receptive fields (LFBs), while an enhanced multi-scale weighting mechanism (AWMS) improves handling of intricate textural patterns. The framework further incorporates: (1) a hybrid optimization objective combining multiple loss terms, and (2) an adaptive training protocol, collectively enhancing reconstruction fidelity. Experimental validation demonstrates consistent superiority over baseline AWSRN across standard benchmarks (B100/Urban100), with measurable gains in both PSNR and structural similarity metrics, confirming the efficacy of our design improvements.

II. Image Super-Resolution Reconstruction Model Based On Awsrn

The core competitiveness of AWSRN lies in its unique network architecture and module design, which efficiently achieves image super-resolution reconstruction. The AWSRN network architecture diagram is shown in Figure 1, which includes two core modules: Local Fusion Block (LFB) and Adaptive Weighted Multi Scale (AWMS) Module.

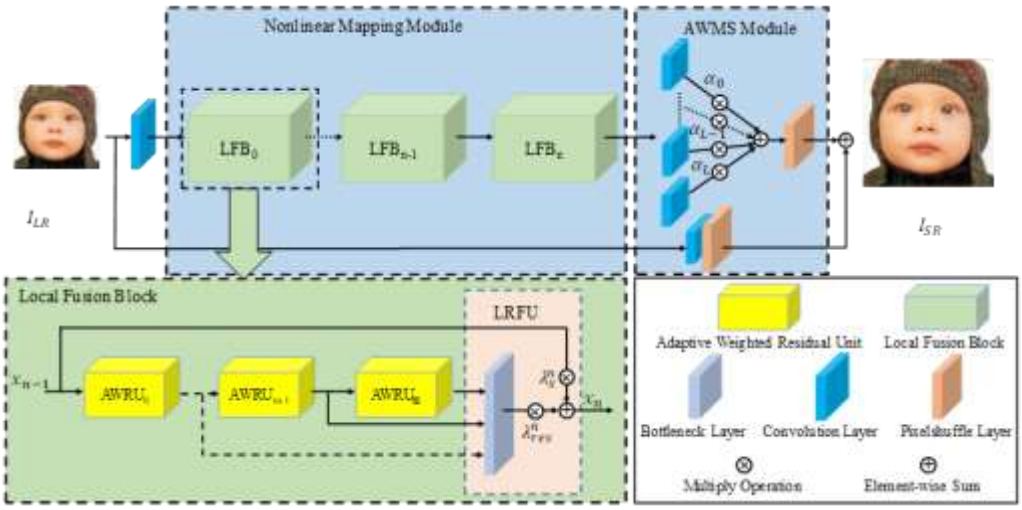

Figure 1. AWSRN Network Architecture Diagram

The LFB module is ingeniously integrated from stacked Adaptive Weighted Residual Units (AWRU) and Local Residual Fusion Units (LRFU). The AWRU unit introduces an adaptive weighting mechanism that can dynamically adjust weights based on the importance of different features, directing the model's attention toward salient features that substantially enhance reconstruction outcomes. This adaptive weighting strategy enhances the efficiency of information and gradient flow without introducing additional parameters, achieving precise learning of image residual information. The LRFU unit further leverages the advantages of local residual fusion, effectively integrating residual information from different AWRU units and enhancing the network's expressive power.

The AWMS module has the ability to fully explore feature information in the reconstruction layer. This module embeds multiple convolutional branches at different scales, which can capture detailed information at different scales in the image. By utilizing convolution operations at multiple different scales, the AWMS module adaptively adjusts the weights of each branch, effectively reducing redundant calculations while ensuring performance. In addition, the AWMS module will intelligently remove redundant scale branches based on the evaluation of network contributions using adaptive weights, and only retain branches that significantly contribute to the reconstruction results. This adaptive weighted multi-scale structure not only improves the efficiency of the network in utilizing feature information, but also significantly enhances the reconstruction performance.

III. Optimization and Improvement Based On AWSRN

Through in-depth analysis of the network structure, reconstruction method, loss function, and training strategy of AWSRN, we propose a series of innovative improvement measures. These architectural refinements simultaneously augment the network's detail preservation capacity while elevating visual quality metrics including sharpness, structural consistency, and photorealistic fidelity.

A. Improvement of Network Structure

The AWSRN network architecture diagram is shown in Figure 2. Firstly, for the optimization of Local Feature Blocks (LFBs), Our optimization efforts primarily targeted the enhancement of Adaptive Weighted Residual Units (AWRUs) and Local Residual Fusion Units (LRFUs) performance characteristics. At the AWRU level, an innovative global context aware weight learning mechanism has been introduced. Extracting global contextual features through Global Average Pooling (GAP), as shown in formula (1).

$$g = \frac{1}{W \times H} \sum_{i=1}^{H} \sum_{j=1}^{W} X(i, j) \quad (1)$$

In this context, $X \in R^{H \times W \times C}$ represents the input feature map, where H and W denote the spatial dimensions of height and width, while the value C corresponds to the total number of channels. $g \in R^{1 \times 1 \times C}$ is the global contextual feature vector. This mechanism not only considers the importance of local features, but also combines global contextual information, by using a two-layer fully connected network (FC) to learn channel attention weights, adaptive weight learning is achieved as shown in formula (2).

$$\alpha = \sigma(W_2 \cdot \delta(W_1 \cdot g + b_1) + b_2) \quad (2)$$

Within this framework, $W_1 \in R^{C/r \times C'}$ and $W_2 \in R^{C' \times C/r}$ denote adjustable weight parameters optimized during training. r represents the compression factor, fixed at 16, while b_1 and b_2 correspond to bias components. The nonlinear activation operator δ is implemented using ReLU, used to normalize weights to the range of [0,1], and $\alpha \in R^{1 \times 1 \times C}$ is the learned channel attention weight. This mechanism can achieve more accurate weight allocation. This improvement significantly enhances AWRU's ability to capture image detail information, providing richer and more accurate feature representations for subsequent image reconstruction.

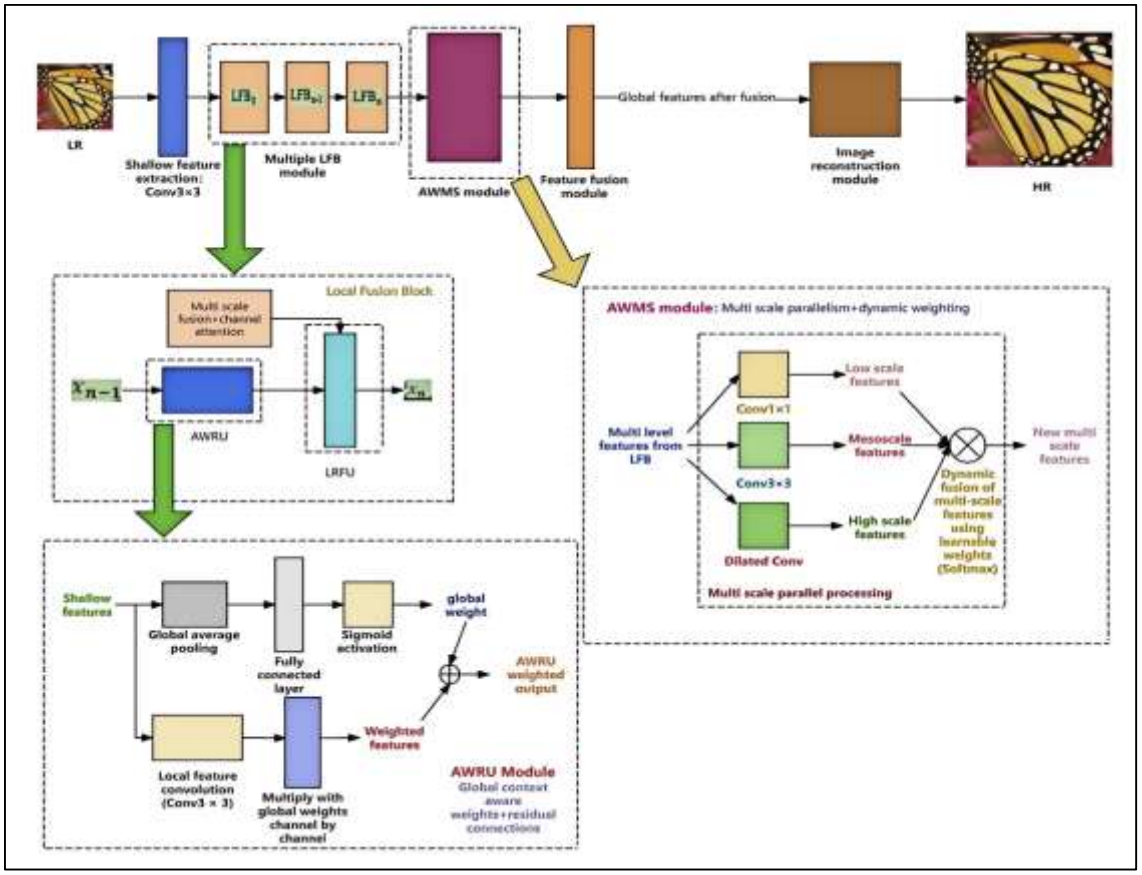

Figure 2. Improved AWSRN network architecture diagram

Secondly, at the LRFU level, we adopted a more optimized feature fusion strategy. Through the incorporation of hierarchical feature integration and focus weighting modules, effective exploitation of multi-level feature representations is accomplished， thereby further improving the quality of reconstructed images. This improvement enables LRFU to more effectively integrate features from different scales and levels, providing more comprehensive and refined feature support for image reconstruction.

In addition, we also optimized the Adaptive Weighted Multiscale (AWMS) module. By introducing richer scale information and more efficient feature extraction strategies, the adaptability of the AWMS module to different scale image features has been significantly improved. This improvement not only enhances the feature extraction capability of the AWMS module, but also increases its sensitivity to image details, thereby further improving the overall performance of AWSRN.

B. Improvement of Reconstruction Methods

Given the challenges posed by diverse image textures, intricate edge details, and noise artifacts, our proposed framework (Figure 3) introduces a novel super-resolution pipeline designed to optimize feature capture completeness and fusion accuracy, thereby advancing reconstruction fidelity.

Our architecture's feature representation phase incorporates a sophisticated unit merging spatially-efficient convolutional decomposition and dynamic feature recalibration. The decomposed convolution process involves: (1) independent spatial filtering per channel, and (2) linear channel combination. The formula for deep convolution is shown in formula (3).

Input feature map $X \in R^{H\times W\times C}$,deep convolution kernel $K \in R^{K\times K\times C}$, output feature map $Y_{depth} \in R^{H\times W\times C}$.

$$Y_{depth}(i,j,c)=\sum_{m=0}^{k-1}\sum_{n=0}^{k-1}K(m,n,c)\cdot X(i+m,j+n,c) \quad (3)$$

Among them, (i, j) is the spatial position, and c is the channel index. The pointwise convolution is shown in formula (4).

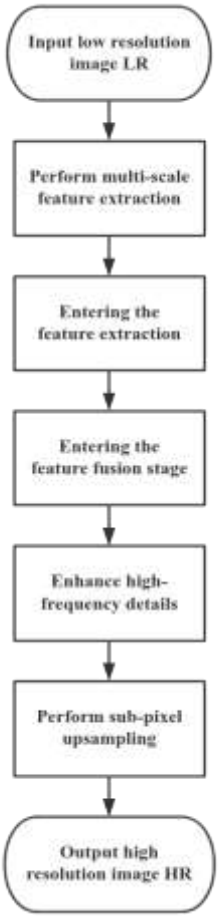

Figure 3. Innovative image super-resolution reconstruction process

Among them, (i, j) is the spatial position, and c is the channel index. The pointwise convolution is shown in formula (4).

Point by point convolution kernel $W \in R^{1\times 1\times C\times C'}$, output feature map $Y_{point} \in R^{H\times W\times C'}$.

$$Y_{point}(i,j,c')=\sum_{C=0}^{C-1}W(1,1,c,c')\cdot Y_{depth}(i,j,c) \quad (4)$$

Among them, c' is the output channel index.

This module not only efficiently extracts multi-scale features from input low resolution images, but also adaptively weights the feature maps

through attention mechanisms. The attention module computes importance scores $A \in R^{H \times W \times C'}$ to dynamically adjust feature representations. These attention coefficients are derived using Equation (5).

Input feature map Y_{point} and generate attention weight A through fully connected or convolutional layers.

$$A = \sigma(W_a \cdot Y_{point} + b_a) \quad (5)$$

Among them, W_a and b_a are learnable parameters, where σ represents the nonlinear activation operation.

The weighted combination of feature representations follows the derivation in formula (6).

Weighted feature map $Y_{att} \in R^{H \times W \times C'}$.

$$Y_{att}(i, j, c') = A(i, j, c') \cdot Y_{point}(i, j, c') \quad (6)$$

By introducing this module, the sensitivity of the model to key image information has been enhanced. The integration of deep features with attention-based weighting enhances the effectiveness of feature learning, but also ensures the richness and accuracy of feature representation.

In the feature fusion stage, we designed a refined fusion strategy based on Long-range structural recurrence and channel dependence. Long-range structural recurrence is used to capture long-range dependencies between feature maps. This strategy calculates the similarity between feature maps, and the calculation process is as follows:

For an input feature representation $X \in R^{C \times H \times W}$ (channel depth C, spatial size $H \times W$), the long-range dependency operations are formulated in (7).

$$Y_i = \frac{1}{C(X)} \sum_{\forall j} f(X_i, X_j) g(X_j) \quad (7)$$

Among them, Y_i is the i-th position of the output feature map. $f(X_i, X_j)$ is a similarity function as shown in formula (8), usually using Gaussian function or dot product to calculate the similarity between positions i and j.

$$f(X_i, X_j) = e^{\theta(X_i)^T \phi(X_j)} \quad (8)$$

Among them, θ and ϕ are linear transformations, usually implemented through 1x1 convolution. $g(Xj)$ is a characteristic transformation function as shown in formula (9), usually also a linear transformation.

$$g(X_j) = W_g X_j \quad (9)$$

$C(X)$ is the normalization factor as shown in formula (10).

$$C(X) = \sum_{\forall j} f(X_i, X_j) \quad (10)$$

Channel dependence is used to dynamically adjust the weights of feature maps to enhance the ability to capture high-frequency details. Given the feature map $Y \in R^{C \times H \times W}$, channel dependence can be achieved through the following steps.

Firstly, calculate the channel attention weight $A \in R^{C \times 1 \times 1}$ as shown in formula (11).

$$A = \sigma(W_2 \delta(W_1 GAP(Y))) \quad (11)$$

Among them, $GAP(Y)$ is a global average pooling operation that compresses the spatial dimension of the features to 1×1 spatial resolution, The learnable parameters W_1 and W_2 correspond to the linear transformation weights, while δ denotes the element-wise activation operator. The sigmoid function σ ensures attention coefficients fall within the unit interval. These weights are then used to recalibrate channel features as mathematically defined in (12).

$$Z = A \otimes Y \quad (12)$$

Where $\otimes$ represents channel wise multiplication operation.

The proposed component facilitates cross-spatial feature synthesis, establishing robust connections between distant image regions. Through channel-wise importance modulation, the framework demonstrates enhanced sensitivity to texture details with improved noise suppression. These refinements yield reconstructions with superior definition and more natural visual continuity.

To rigorously validate our approach, we performed extensive training iterations on the DIV2K benchmark, adhering to established super-resolution evaluation protocols. Quantitative comparisons with current AWSRN architectures reveal that our novel feature integration framework delivers superior perceptual quality. The method exhibits particular advantages in processing challenging visual patterns containing intricate structures and sharp transitions.

C. Improvement of Loss Function

The conventional loss formulation adopted from AWSRN studies incorporates both MSE and PSNR-optimized components. While this framework provides basic pixel-level fidelity measurement between reconstructed and reference images, it demonstrates notable limitations in preserving fine structural details, textural patterns, and perceptual authenticity. These traditional loss functions often result in reconstructed images being too smooth, lacking realistic texture details and sharp edges.

To address these limitations, we enhance the AWSRN optimization framework through a hybrid loss formulation integrating perceptual and adversarial components. The perceptual term quantifies feature-level discrepancies in texture, edge, and structural patterns by evaluating VGG-encoded representations (using a pre-trained VGG network in our implementation). Simultaneously, the adversarial component employs GAN-based discriminative evaluation to improve visual authenticity and realism in reconstructions [7].

The perceptual discrepancy metric computes either L1 or L2 norms between the feature activations of reconstructed and reference images, extracted from designated VGG network layers. This formulation, mathematically expressed in Equation (13), serves to reduce semantic-level feature distortions in the output.

$$L\_perceptual=\sum\_l(1/N\_l)*\square\varphi\_l(I\_generated)-\varphi\_l(I\_target)\square\hat{}2 \quad (13)$$

Among them, $L\_perceptual$ represents the perceptual loss, φ_l represents the features extracted by the pre trained network at the l-th layer, $I2$ generated represents the generated image (i.e. reconstructed image), $I2$ target represents the target image (i.e. original image), N1 represents the dimension of the l-th layer features, $\square\hat{}2$ represents the L2 norm (i.e. the square of Euclidean distance).

The computation involves aggregating L2-norm distances across all feature map layers to derive the cumulative perceptual discrepancy metric. This quantitative measure evaluates the divergence between reconstructed and reference images within deep feature representations.

The adversarial optimization framework employs a discriminative network trained to differentiate super-resolved outputs from ground truth samples, while simultaneously optimizing the generator (AWSRN architecture) to produce visually plausible results capable of bypassing this discrimination, as mathematically formulated in Eq. (14).

$$L\_adversarial=-E[\log(D(G(z)))] \quad (14)$$

Among them, $L\_adversarial$ represents adversarial loss, D represents discriminator, G represents generator (i.e. AWSRN), z represents random noise or low resolution image input to the generator, $E[...]$ represents expectation.

The proposed optimization framework combines perceptual and adversarial losses through weighted summation, yielding the final objective function as defined in Equation (15).

$$L\_total=\lambda\_perceptual*L\_perceptual+\lambda\_adversarial*L\_adversarial \quad (15)$$

Among them, L_total represents the total loss function, while $\lambda\_perceptual$ and $\lambda\_adversarial$ represent the weight coefficients of perceptual loss and adversarial loss, respectively.

D. Training Strategy Improvement

To address AWSRN's training challenges including slow convergence, local optimum trapping, and inadequate high-frequency feature learning, we develop a Multi-Phase Training Scheme (MPTS). This framework implements: (1) A hierarchical curriculum learning approach that incrementally processes images from reduced to full resolution, enhancing detail learning [8]; (2) An adaptive learning rate mechanism incorporating cosine annealing with warm restarts for optimized convergence; (3) A Feature-Adaptive Sample Selection (FASS) module that prioritizes high-information-content samples based on feature distribution analysis [9].

IV. Experiment and Analysis

A. Train Data Set

The DIV2K benchmark comprises 800 training and 100 validation images in high-resolution (HR) format, all exhibiting exceptional visual quality with well-preserved fine details. This collection serves as an excellent resource for developing and testing super-resolution methods. A key advantage of this dataset is its systematic degradation pipeline, which allows generation of low-resolution (LR) counterparts at multiple magnification levels (×2, ×3, ×4) through automated scripts. This standardized preprocessing ensures dataset uniformity while simplifying experimental setup [10].

Furthermore, DIV2K incorporates both bicubic interpolation and configurable degradation models to generate more authentic low-resolution counterparts, enabling comprehensive evaluation of SR methods. The dataset's premium-quality images serve as an optimal training basis for AWSRN, with diverse samples enhancing the network's adaptability and reconstruction quality. The validation subset facilitates rigorous assessment of model precision and stability, verifying practical deployment readiness.

The DIV2K dataset's superior image quality establishes an optimal training basis for AWSRN, with its diverse samples significantly enhancing the network's cross-domain adaptability and reconstruction fidelity. For performance verification, the dedicated validation set enables comprehensive assessment of the model's precision and stability, confirming its operational effectiveness in real-world scenarios [11].

B. Experimental Hardware Configuration

The hardware configuration of this experiment adopts AMAX workstation, equipped with Intel Xeon Gold 6254 processor (18 cores, 36 threads, main frequency 3.1GHz) and 32GB DDR4 2666MHz ECC memory, ensuring high-performance computing and data reliability. The operating system is Ubuntu 18.04 LTS, and the graphics card is NVIDIA GeForce RTX 2080 Ti (11GB GDDR6 VRAM), supporting CUDA and cuDNN acceleration, suitable for training deep learning models. The storage is configured as a 1TB NVMe SSD for fast reading and writing of datasets and model files.

C. Experimental Process

To validate our optimization approach for super-resolution generation, we first trained the model using DIV2K data. For quantitative evaluation, benchmark datasets B100 and Urban100 were employed, with reconstruction quality assessed through two established metrics: PSNR (quantifying pixel-level accuracy) and SSIM (measuring structural preservation). These measurements enable systematic comparison between generated and reference high-resolution images, providing objective performance evaluation. The experimental procedure consists of:

a). Data curation involves systematically pairing high-resolution source images with their synthetically degraded counterparts (generated through resolution reduction) to create organized training and evaluation subsets.

b). The experimental framework involves implementing an adaptive-weighted SR network in Python, incorporating structural refinements to the Local Fusion Block (LFB). Key enhancements include optimizing the Adaptive Weighted

Residual Unit (AWRU) and Local Residual Fusion Unit (LRFU), along with improvements to the Adaptive Weighted Multi-Scale (AWMS) module, all trained using the Adam optimizer.

c). Data preprocessing: Preprocessing the selected image data, including image normalization, cropping, scaling, and other operations, in order to input it into the model for training and testing.

d). The training phase involves feeding low-resolution (LR) samples from the benchmark dataset into the network, with corresponding high-resolution (HR) images serving as ground truth targets for supervised learning.

e). During the assessment phase, the optimized network processes low-resolution test samples to perform super-resolution restoration. Comparative visual results in Figure 4 demonstrate enhanced detail preservation, with panel (a) displaying baseline AWSRN outputs and panel (b) showing our improved reconstruction quality.

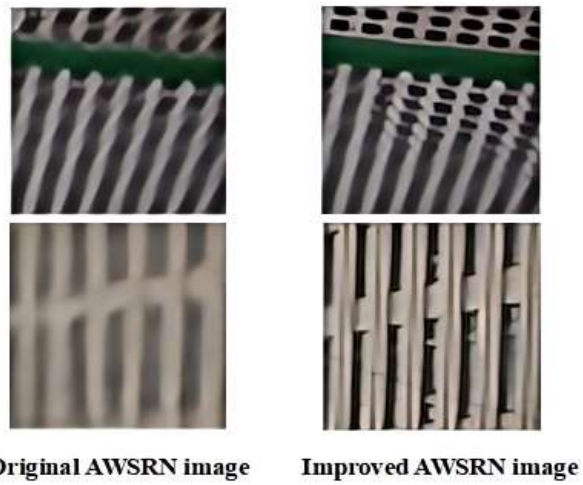

Figure 4. Comparison of image reconstruction details

To quantitatively evaluate reconstruction quality, we employ two established metrics: PSNR measures pixel-level fidelity, while SSIM assesses structural preservation relative to ground truth HR references. Quantitative comparisons at ×4 magnification appear in Table I, with corresponding ×8 results presented in Table II.

TABLE I. QUANTITATIVE COMPARISON ON A SCALE FACTOR OF 4

Scale	Model	B100 PSNR/SSIM	Urban100 PSNR/SSIM
4	AWSRN	27.64/0.7385	26.29/0.7930
	Improved AWSRN	28.47/0.7592	27.35/0.8169

TABLE II. QUANTITATIVE COMPARISON ON A SCALE FACTOR OF 8

Scale	Model	B100 PSNR/SSIM	Urban100 PSNR/SSIM
8	AWSRN	24.80/0.5967	22.45/0.6174
	Improved AWSRN	25.32/0.6214	23.18/0.6438

D. Experimental Results

Quantitative analysis reveals consistent performance gains across all test conditions. For 4× super-resolution, the enhanced ASWRN architecture demonstrates measurable improvements over baseline AWSRN, with B100 dataset showing PSNR/SSIM gains of +0.83 dB/+0.0207 and Urban100 achieving +1.06 dB/+0.0239 improvements. At 8× magnification, quality metrics further improve, with B100 registering +0.52 dB/+0.0247 and Urban100 showing +0.73 dB/+0.0264 enhancements. These progressive gains with increasing scale factors confirm the optimization's effectiveness in bridging the gap between reconstructed and reference images.

The empirical analysis confirms the enhanced AWSRN framework's efficacy in single-image super-resolution applications, demonstrating substantial improvements in both computational efficiency and reconstruction fidelity compared to existing approaches. This optimized architecture achieves superior high-frequency detail recovery and perceptual quality, offering valuable insights for advancing SR algorithm development and computer vision applications [12].

V. CONCLUSIONS

This study proposes an improved AWSRN super-resolution network algorithm that effectively addresses the efficiency and optimization issues in

image super-resolution reconstruction. By optimizing the LFB, LRFU, and AWMS modules, the ability to capture details and reconstruction results have been significantly improved. Empirical evaluations on B100 and Urban100 benchmarks demonstrate consistent metric improvements (PSNR/SSIM) in the enhanced model, with reconstructions exhibiting superior fidelity to ground truth references. This framework introduces novel paradigms for single-image super-resolution while advancing computer vision methodologies. Future directions include: (1) architectural refinements for scenario-specific performance tuning, (2) development of robust optimization protocols to enhance model stability and computational efficiency. The current module-level optimizations establish a strong foundation for subsequent algorithmic developments.

REFERENCES

[1] Dong, C., Loy, C. C., & Tang, X, ″Accelerating the Super-Resolution Convolutional Neural Network,″ in Proc. European Conference on Computer Vision (ECCV), Springer, pp. 391–407, 2020.

[2] J. Kim, J. K. Lee, and K. M. Lee, ″Deep Recursive Residual Network for Image Super-Resolution,″ in Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3566–3575, 2021.

[3] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, ″ Residual Recursive Network for Image Super-Resolution,″ in Proc. AAAI Conference on Artificial Intelligence (AAAI), vol. 36, no. 3, pp. 3456−3464, 2022.

[4] Wang, X., Yu, K., Dong, C., & Loy, C. C. (2021). ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 63-79). Springer.

[5] C. Wang, Z. Li, and J. Shi, ″Lightweight Image Super-Resolution with Adaptive Weighted Learning Network,″ arXiv preprint arXiv:1904.02358, 2019.

[6] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, ″ Residual Feature Aggregation Network for Image Super-Resolution,″ in Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2358-2367,2022.

[7] X. Wang, K. Yu, C. Dong, and C. C. Loy, ″Recovering Realistic Texture in Image Super-Resolution by Deep Spatial Feature Transform,″ in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 606–615, 2018.

[8] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, ″Image Super-Resolution Using Very Deep Residual C-hannel Attention Networks,″ IEEE Transactions on Patt ern Analysis and Machine Intelligence, vol. 43, no. 10, pp.3551-3567,2021.

[9] X. Chen, J. Wang, and Y. Guo, ″Dynamic Learning Rate Scheduling for Deep Neural Networks with Cosine Annealing and Warm Restarts, ″ Neural Networks, vol. 145, pp. 1–12, 2022.

[10] Z. Wang, J. Chen, and S. C. H. Hoi, ″Deep Learning for Image Super-Resolution: A Survey,″ IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 10, pp. 3365–3387, 2021.

[11] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, “Image Super-Resolution Using Very Deep Residual C hannel Attention Networks,” IEEE Transactions on P-at tern Analysis and Machine Intelligence, vol. 43, no. 10, pp. 3551–3567, 2021.

Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, “Residual Dense Network for Image Super-Resolution,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 5, pp. 2480–2495, 2022.