

Channel Distortion-Aware Deep Image Super-Resolution Reconstruction for Wireless Internet of Things Applications



Hongbo Yu^{*}, Shurui Wang, Chao Liu, Zhan Jin

School of Communication and Electronic Engineering, Qiqihar University, Qiqihar 161006, China

Corresponding Author Email: xgk029yhb@126.com

Copyright: ©2026 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.430202>

ABSTRACT

Received: 6 October 2025

Revised: 20 February 2026

Accepted: 10 March 2026

Available online: 30 April 2026

Keywords:

Wireless Internet of Things, image super-resolution, channel distortion modeling, end-to-end learning, dynamic feature modulation, channel consistency loss, lightweight deployment, edge computing

Wireless Internet of Things (WIoT) visual applications have been widely adopted in areas such as intelligent surveillance, agricultural remote sensing, and unmanned aerial vehicle (UAV) inspection. As a key technique for enhancing low-resolution images, image super-resolution plays a critical role in enabling these applications. However, in WIoT scenarios, resource-constrained sensing devices often capture low-resolution images. Moreover, image transmission over wireless channels is susceptible to various distortions, including multipath fading, additive noise, and packet loss, which severely degrade image quality. Conventional image super-resolution methods typically adopt a two-stage pipeline that performs distortion removal followed by super-resolution. Such a decoupled strategy overlooks the intrinsic coupling between channel distortion and SR reconstruction, making it difficult to achieve joint optimization. In addition, existing SR networks generally lack adaptability to dynamic channel conditions and often involve large model sizes, which limits their deployment on edge devices with strict constraints on computational resources and real-time performance. To address these challenges, this paper proposes an end-to-end deep image super-resolution reconstruction method that incorporates channel distortion modeling. The proposed approach consists of three key components: (1) a differentiable channel distortion module (DCDM) for accurate modeling and learnable adaptation of channel effects; (2) a channel-aware dynamic modulation mechanism within the SR network to enhance adaptability to varying channel conditions; and (3) a channel-consistency loss to jointly optimize distortion removal, detail preservation, and perceptual quality. The proposed method can be efficiently deployed on WIoT edge devices, providing a reliable image reconstruction solution for low-bandwidth visual applications. It demonstrates significant theoretical value and practical potential.

1. INTRODUCTION

With the rapid iteration of Wireless Internet of Things (WIoT) technology, visual applications have been widely penetrated into multiple key fields such as intelligent surveillance, agricultural remote sensing imaging, and unmanned aerial vehicle (UAV) inspection [1-3]. Such applications continuously increase the demand for image quality. As a core technique for optimizing low-resolution image quality, image super-resolution technology has become a key support for the large-scale deployment of WIoT visual applications [4-5]. However, image transmission in WIoT scenarios faces dual inherent bottlenecks: on the one hand, terminal sensors are constrained by size [6], power consumption, and cost [7], resulting in inherently low-resolution captured images [8] with insufficient detail information [9]; on the other hand, the transmission characteristics of wireless channels easily introduce multiple distortions such as multipath fading [10], additive noise [11], and packet loss [12], which further aggravate image quality degradation and seriously affect the performance of downstream tasks such as object detection and image

recognition.

At present, image super-resolution research for WIoT scenarios still has four core defects, which restrict the practical deployment and application of the technology. Traditional methods generally adopt a decoupled processing strategy of first distortion removal and then super-resolution, which separates the intrinsic coupling relationship between channel distortion and super-resolution tasks, leading to the loss of useful texture information during distortion removal, and making it difficult to accurately recover details in the super-resolution stage, resulting in reconstructed images prone to artifacts and edge blurring [13-15]. Existing super-resolution networks are mostly designed for general purposes and do not fully consider the dynamic variation characteristics of WIoT channels. They lack adaptive capability for different intensities of channel impairments, have insufficient robustness, and are difficult to cope with complex and dynamic wireless transmission environments [16, 17]. Meanwhile, existing loss function designs only focus on optimizing super-resolution reconstruction quality and lack dedicated constraints for channel distortion. Excessive smoothing during reconstruction easily causes the loss of key

texture information in images and cannot guarantee the realism of reconstruction results [18, 19]. In addition, existing high-performance super-resolution models have a large number of parameters and high computational complexity, which makes it difficult to adapt to the limited computational resources and power constraints of WIoT edge devices, and cannot meet real-time processing requirements [20, 21].

The research objective of this paper is to break through the limitations of existing methods and propose an end-to-end deep image super-resolution method incorporating channel distortion modeling, to achieve collaborative optimization of distortion removal and super-resolution in WIoT scenarios, while taking into account both high reconstruction quality and lightweight, real-time deployment requirements on edge devices, thereby providing a reliable solution for low-bandwidth IoT visual applications. The main contributions are as follows:

- A differentiable channel distortion module (DCDM) is proposed, which breaks through the limitations of traditional fixed channel simulation, and unifies multiplicative fading, additive noise, and random block loss into differentiable operations. With learnable parameters and end-to-end joint training with the super-resolution network, an adversarial adaptive sample generation mechanism is formed.
- A Channel-Aware Super-Resolution Network (CASR-Net) is designed. Based on an improved residual channel attention network, a channel state feature extractor and a dynamic feature modulation layer are introduced to solve the problem that traditional networks cannot adapt to dynamic channels.
- A joint loss function with channel consistency constraints is constructed. A channel consistency loss is introduced to preserve the reversibility of channel distortion. The L1 norm is used to constrain the pixel consistency between the re-distorted image and the original distorted image to avoid texture loss. Combined with Charbonnier loss, edge loss, and Visual Geometry Group (VGG)-19 perceptual loss, collaborative optimization of distortion removal, detail preservation, and visual realism is achieved, improving both subjective and objective reconstruction quality.
- A lightweight deployment scheme for edge devices is proposed, integrating knowledge distillation and structured pruning to solve the problems of large model size and slow inference.

The remainder of this paper is organized as follows: Section 2 reviews related research and analyzes the current status and deficiencies in image super-resolution under wireless channel distortion and conditional super-resolution networks; Section 3 describes the proposed method in detail; Section 4 verifies the effectiveness and superiority of the method through multiple experiments; Section 5 discusses the experimental value, method limitations, and future improvement directions; Section 6 summarizes the main work and conclusions, and clarifies the theoretical and practical value of this study.

2. PROPOSED METHOD

2.1 Overall framework of the method

The deep image super-resolution reconstruction method incorporating channel distortion modeling proposed in this

paper adopts an end-to-end collaborative optimization architecture as a whole. The core logic lies in realizing the deep coupling between channel distortion modeling and image super-resolution reconstruction, breaking the limitations of traditional decoupled processing. The overall process can be described as follows: taking the clean low-resolution image I_{LR} as the initial input, it is first fed into the DCDM. This module completes the joint modeling of multipath fading, additive noise, and random block loss based on the learnable parameter set θ_c , and generates a distorted low-resolution image that matches the characteristics of real wireless channels:

$$I_{LR}^{dist} = D_{\theta_c}(I_{LR}) \quad (1)$$

Then, I_{LR}^{dist} is input into the CASR-Net. Through channel state feature extraction, dynamic feature modulation, and collaborative attention mechanism, the network completes accurate restoration of distorted features and super-resolution reconstruction, and outputs the high-resolution reconstructed image:

$$I_{SR} = G(I_{LR}^{dist}, \theta_g) \quad (2)$$

where, θ_g denotes the parameters of the super-resolution network. Finally, the reconstructed result is supervised by a joint loss function:

$$L_{total} = L_{SR} + \lambda_1 L_{dist_consistency} + \lambda_2 L_{perceptual} \quad (3)$$

The gradients are backpropagated along the super-resolution network to the channel distortion module, realizing the adaptive collaborative update of θ_c and θ_g , ensuring that the channel distortion modeling and super-resolution reconstruction processes are mutually adapted and collaboratively optimized, thereby fundamentally improving the super-resolution reconstruction performance of distorted images in WIoT scenarios. The overall framework is shown in Figure 1.

2.2 Differentiable channel distortion module

Multipath fading is one of the core impairments in WIoT channels that leads to image quality degradation. Its frequency-selective attenuation characteristic destroys the distribution of image frequency-domain features. Traditional fixed channel simulation methods are difficult to adapt to dynamically changing channel environments, and cannot achieve collaborative optimization with the super-resolution network. The multipath fading simulation mechanism designed in this paper has the core innovation of realizing differentiable modeling and parameter-adaptive learning of fading characteristics, without manually presetting fixed channel parameters, and can directly optimize fading features from distorted images in a backward manner. The internal structure and processing flow of DCDM are shown in Figure 2. Specifically, first the input clean low-resolution image I_{LR} is subjected to a two-dimensional Fourier transform, converting the image from spatial domain to frequency domain, and obtaining frequency-domain features:

$$F(I_{LR}) = FFT2(I_{LR}) \quad (4)$$

where, $FFT2$ denotes the two-dimensional fast Fourier

transform operation. Then, a complex frequency response $H(u, v)$ is introduced to modulate the frequency-domain features to realize frequency-selective attenuation simulation. The magnitude of $H(u, v)$ follows a Rayleigh or Rician distribution, and the phase follows a uniform distribution in $[0, 2\pi)$. Its magnitude and phase parameters are all learnable variables, forming differentiable channel fading coefficients. The mathematical expression of the frequency-domain modulation process is:

$$F_{fade} = F(I_{LR}) \odot H(u, v) \quad (5)$$

where, \odot represents element-wise Hadamard product.

Finally, the frequency-domain features are transformed back to the spatial domain through the two-dimensional inverse Fourier transform to obtain the image with multipath fading distortion:

$$I_{fade} = IFFT2(F_{fade}) \quad (6)$$

The core advantage of this design is that all parameters of $H(u, v)$ can be adaptively updated through backpropagation, which can accurately match the multipath fading intensity of different wireless channels, and realize end-to-end collaborative optimization between fading characteristics and super-resolution reconstruction.

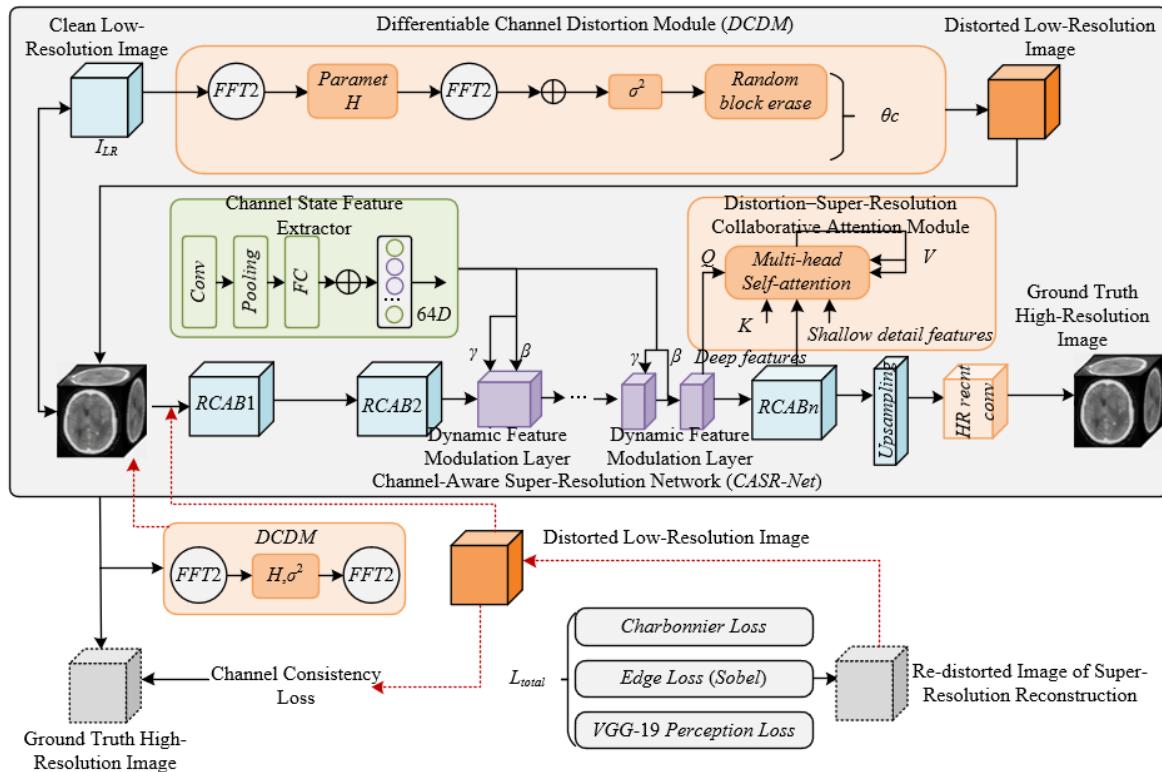


Figure 1. Overall framework of the proposed end-to-end deep image super-resolution reconstruction method

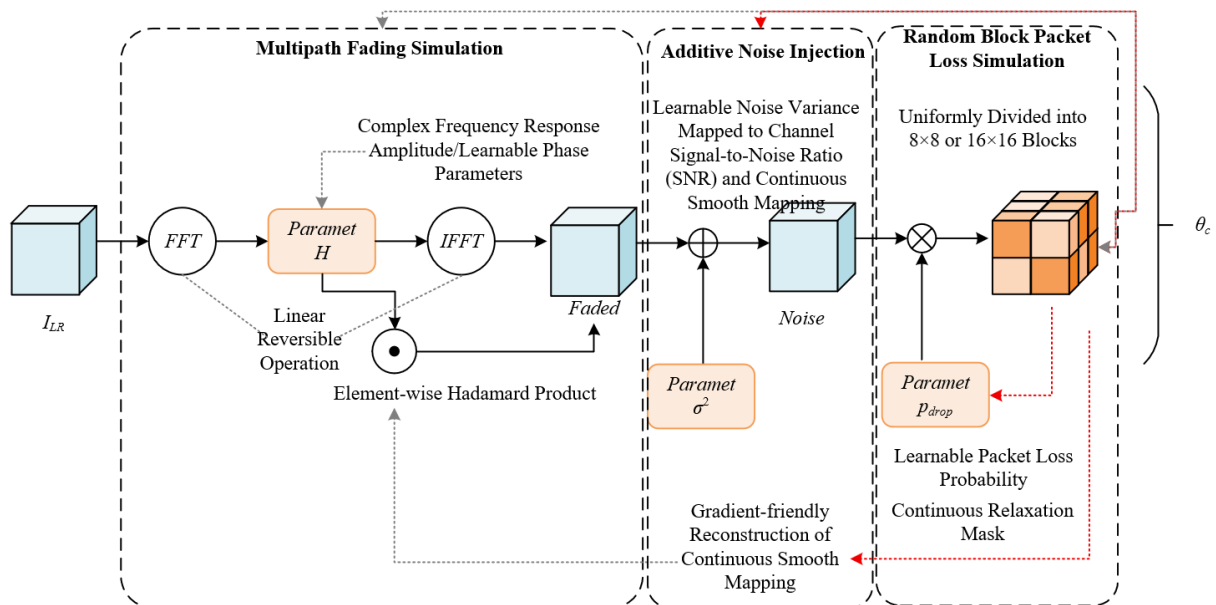


Figure 2. Internal structure and processing flow of the differentiable channel distortion module (DCDM)

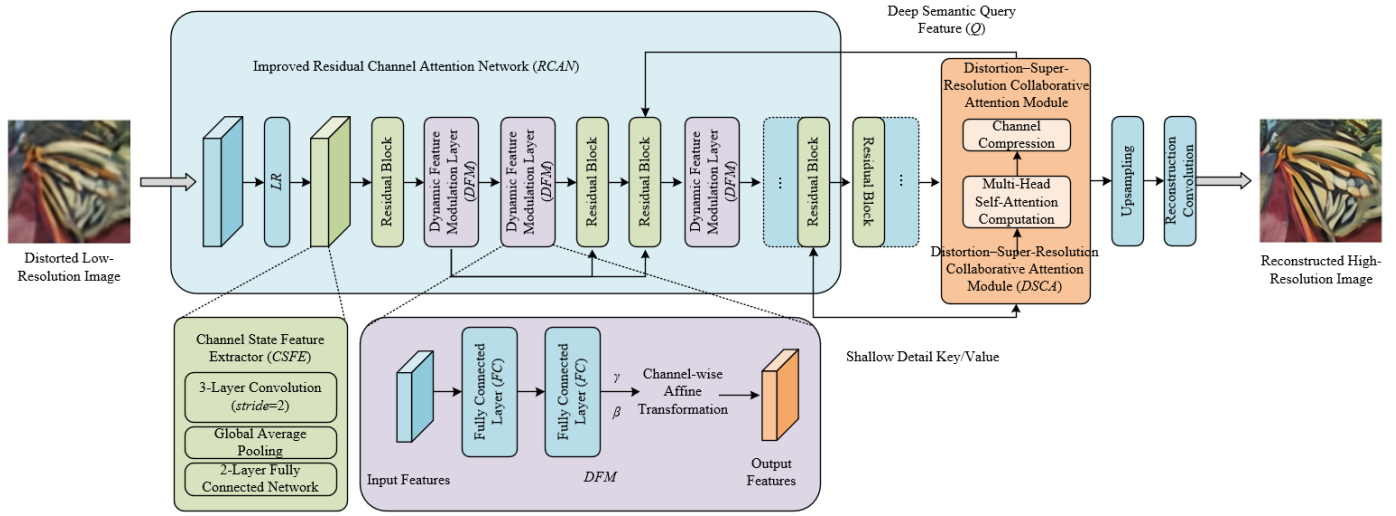


Figure 3. Architecture of the Channel-Aware Super-Resolution Network (CASR-Net)

Additive noise is another typical channel impairment in wireless transmission, and its intensity is directly related to the signal-to-noise ratio (SNR) of the channel. Traditional simulation methods with fixed noise variance cannot adapt to dynamically changing channel noise levels, which easily leads to insufficient noise suppression ability of the super-resolution network. The additive noise superposition mechanism proposed in this paper realizes differentiable adaptive adjustment of noise intensity. The core idea is to treat the noise variance σ_n^2 as a learnable parameter and include it into the module parameter set θ_c , and jointly train it with the super-resolution network. Specifically, taking the image I_{noise} after multipath fading distortion as input, zero-mean Gaussian white noise n is added, where the noise satisfies $n \sim N(0, \sigma_n^2)$. The mathematical expression of the superposition process is:

$$I_{noise} = I_{fade} + n \quad (7)$$

where, the noise variance σ_n^2 is obtained through a lightweight fully connected network mapping and is dynamically associated with the channel SNR. When the channel SNR decreases, the network automatically increases σ_n^2 to simulate a strong noise environment; when the channel SNR increases, it automatically decreases σ_n^2 to match a weak noise transmission scenario. Different from traditional fixed noise simulation, this differentiable noise superposition mechanism allows the network to autonomously learn the distribution characteristics of different channel noises. By optimizing σ_n^2 through backpropagation, the noise simulation is more consistent with real wireless channel environments, while ensuring the differentiability of the noise addition operation, providing a foundation for subsequent end-to-end joint training with the super-resolution network.

The collaborative effect of multipath fading simulation and additive noise superposition constitutes the core distortion modeling capability of the DCDM module. Both follow a differentiable design principle, and all parameters can be updated through backpropagation. This design breaks the limitation of traditional separation between channel distortion simulation and super-resolution reconstruction, enabling channel distortion modeling to directly respond to super-resolution reconstruction requirements. Through parameter collaborative optimization, the simulated distortion features

are more consistent with real wireless channel impairments, providing accurate channel state information support for the subsequent CASR-Net dynamic modulation, while avoiding the subjectivity and limitations caused by manually setting channel parameters.

For the burst packet loss phenomenon in WIoT links, this paper designs a structured random block erasure mechanism to simulate real channel packet loss distortion. Different from traditional unstructured global pixel random erasure methods, the image after noise processing is uniformly divided into non-overlapping image blocks of two sizes, 8×8 and 16×16 . The block structure is consistent with wireless communication data packet transmission formats, which can more accurately reproduce real transmission packet loss characteristics. The module independently samples each image block using a learnable packet loss probability p_{drop} , and performs pixel zeroing on blocks where packet loss is triggered. The overall distortion mapping relationship can be expressed as:

$$I_{LR}^{dist} = M \odot I_{noise} \quad (8)$$

where, I_{LR}^{dist} is the intermediate image after multipath fading and additive noise processing, and M is a mask matrix corresponding to the block structure. The mask values are dynamically controlled by the learnable parameter p_{drop} . The final output is the complete distorted low-resolution image I_{LR}^{dist} . The packet loss probability p_{drop} is included in the overall learnable parameter set θ_c of the module and can be adaptively adjusted during training, effectively improving the matching degree between distortion simulation and real WIoT channel environments.

This paper performs gradient-friendly reconstruction of all operations in the DCDM, fundamentally solving the inherent limitation that traditional discrete channel simulation cannot be integrated into end-to-end training. The two-dimensional Fourier transform and inverse transform are linear, invertible, and differentiable operations, and additive Gaussian noise superposition is a continuous and smooth mapping; both can directly support gradient backpropagation. For the inherently non-differentiable discrete hard-threshold block erasure problem, this paper introduces a continuous relaxation mask mechanism to replace the original binary discrete mask, transforming step-wise block erasure into a continuous weighted differentiable operation, so that block packet loss

simulation also satisfies differentiability conditions.

All distortion parameters of the module are uniformly denoted as the parameter set θ_c , and the entire distortion modeling process forms a continuous differentiable mapping:

$$I_{LR}^{dist} = D_{\theta_c}(I_{LR}) \quad (9)$$

Gradients can be fully propagated bidirectionally between the distortion module and the super-resolution network, realizing adaptive collaborative optimization between channel distortion modeling and image super-resolution reconstruction. This fundamentally changes the traditional decoupled processing mode where preprocessing and reconstruction models are independent and parameters are uncoupled, enabling channel distortion characteristics to actively adapt to super-resolution reconstruction requirements, and further improving reconstruction stability under complex channel environments.

2.3 Channel-Aware Super-Resolution Network (CASR-Net)

The CASR-Net proposed in this paper is based on an improved residual channel attention network as the backbone. The network does not perform redundant reconstruction of the basic residual attention structure, but instead introduces two core modules: a channel state feature extractor and a dynamic feature modulation layer. These two modules fundamentally solve the inherent defects of traditional super-resolution networks, which have static feature representations and cannot autonomously perceive dynamic channel impairments in WIoT, and realize implicit channel perception and adaptive super-resolution reconstruction of distorted images.

The architecture of CASR-Net is shown in Figure 3. The channel state feature extractor is the core component for realizing channel perception capability in the network. This module performs progressive spatial downsampling on the input distorted image features through a three-layer convolution structure with stride 2, compressing the feature dimension to 8×8 , and fully aggregating local distortion spatial distribution information. Then, global average pooling is applied to complete global distortion feature aggregation, followed by two consecutive fully connected networks to perform feature dimension normalization and nonlinear mapping, and finally output a 64-dimensional dense channel embedding vector z_{ch} . This vector can fully represent channel impairment information such as multipath fading strength, noise level, and block packet loss degree corresponding to the current image. Different from traditional conditional super-resolution methods relying on manually annotated channel parameters, the channel state feature extractor directly performs implicit channel state encoding from the distorted image, without additional channel supervision labels, and all parameters can be jointly optimized end-to-end with the network backbone.

The dynamic feature modulation layer is uniformly deployed after each residual block in the network and is responsible for adaptive transformation of channel-aware information into feature space. This module takes the embedding vector z_{ch} output by the channel state feature extractor as the only input, and generates channel scaling parameter γ and channel bias parameter β through two independent lightweight fully connected mapping networks $\phi_\gamma(z_{ch})$ and $\phi_\beta(z_{ch})$, respectively. Then, channel-wise affine

transformation is performed on the output features of the residual block. The mathematical expression is:

$$F_{out} = \gamma \odot F_{in} + \beta \quad (10)$$

where, F_{in} is the input feature map before modulation, \odot represents channel-wise Hadamard product, and F_{out} is the output feature map after channel-adaptive modulation.

This dynamic modulation mechanism can autonomously adjust the response strength of each feature channel according to real-time channel impairment levels, and has fine-grained adaptive optimization characteristics. In strong noise channel environments, the network automatically reduces the gain of high-frequency noise-related channels to suppress noise interference expansion in the reconstruction process; in channel environments with severe block packet loss, the network autonomously enhances the weights of texture-related channels in missing regions to strengthen detail restoration capability. Compared with traditional fixed conditional fusion methods such as feature concatenation and element-wise addition, channel-wise dynamic affine modulation achieves fine-grained channel-adaptive feature regulation, and significantly improves the adaptability and robustness of the network to complex dynamic wireless channel environments with only a small increase in computational cost.

The distortion-super-resolution collaborative attention module, as a core enhancement component of CASR-Net for accurate distortion restoration, adopts a multi-head self-attention variant structure. The core innovation is to break the limitation that query, key, and value in traditional attention mechanisms are from the same source, and achieve collaborative optimization of distortion localization and detail restoration through heterogeneous feature interaction. The query vector is taken from deep semantic features of the network. Deep features, after multiple rounds of residual learning and dynamic modulation, have stronger global semantic representation capability and can accurately identify the spatial distribution of severely distorted regions in the image. The key vector and value vector are taken from shallow features. Shallow features retain more detailed texture information of the original distorted image and can provide accurate feature support for distortion region restoration. The module maps deep query features and shallow key-value features into the same feature space through lightweight linear mapping. The attention weight calculation expression is:

$$\text{Attn}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (11)$$

where, Q , K , V are deep query features, shallow key features, and shallow value features respectively, and d_k is the feature dimension, used to alleviate attention weight deviation caused by dimensionality disaster. To control computational complexity, the module adopts a channel compression strategy, reducing the input feature channels to 1/2 of the original before attention computation, and limits the number of attention heads to 4, ensuring that the computational cost of the module increases only by 15% compared with the baseline network, achieving a balance between performance and efficiency.

The core role of the Distortion-Super-resolution Collaborative Attention (DSCA) module is to establish a collaborative relationship between deep semantic distortion localization and shallow detail restoration, guiding network

resources to focus on severely distorted regions. Deep semantic query features can accurately capture the spatial location and intensity of distortions such as noise and packet loss, and through attention weight allocation, focus the restoration on regions with obvious artifacts and missing textures. Shallow key-value features provide original texture details for these regions, and through attention-weighted fusion, shallow effective features are transferred to deep features to guide precise restoration. This design effectively solves the problems of inaccurate restoration of distorted regions and excessive smoothing of non-distorted regions in traditional super-resolution networks. Together with Channel State Feature Extractor (CSFE) for channel state extraction and Dynamic Feature Modulation (DFM) for dynamic feature modulation, it forms a synergistic effect, enabling the network to adaptively allocate restoration resources according to channel impairment distribution, suppress artifacts while maximally preserving original image details, and further improve both subjective and objective quality of reconstructed images.

$$Attn(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

2.4 Joint loss function and training strategy

The core of the joint loss function design is to realize deep synergy between super-resolution reconstruction and channel distortion modeling, breaking the limitation that traditional loss functions only focus on reconstruction quality while ignoring channel characteristics. Through multi-loss collaborative constraints, reconstruction accuracy, detail preservation, and channel adaptability are jointly considered. The total loss expression is:

$$L_{total} = L_{SR} + \lambda_1 L_{dist_consistency} + \lambda_2 L_{perceptual} \quad (12)$$

where, λ_1 and λ_2 are the weight coefficients of channel consistency loss and perceptual loss respectively. Experimental results show that when set to 0.5 and 0.3, optimal balance among losses can be achieved. The innovation of this loss function lies in introducing reversibility constraints of channel distortion into super-resolution training, so that the super-resolution network can not only perform image restoration but also adapt to dynamically changing channel environments, fundamentally solving the insufficient robustness of traditional super-resolution models in complex channel scenarios.

The super-resolution reconstruction loss L_{SR} adopts a combination of Charbonnier loss and edge loss to jointly consider reconstruction accuracy and contour integrity. The Charbonnier loss is a smooth approximation of L1 loss, and its expression is:

$$L_{Charbonnier} = \sqrt{\|I_{SR} - I_{HR}\|^2 + \epsilon^2} \quad (\epsilon = 1e-3) \quad (13)$$

It can effectively alleviate the edge blurring problem caused by traditional L1 loss and reduce the influence of outliers on training. The edge loss is based on the Sobel gradient operator, which strengthens contour restoration by computing the gradient difference between reconstructed image and ground truth high-resolution image. The expression is:

$$L_{edge} = \|Sobel(I_{SR}) - Sobel(I_{HR})\|_1 \quad (14)$$

It makes the edge details of reconstructed images clearer and more three-dimensional. The channel consistency loss $L_{dist_consistency}$ is the core innovative loss, and its expression is:

$$L_{dist_consistency} = \|C_{\theta_c}(I_{SR}) - I_{LR}^{dist}\|_1 \quad (15)$$

where, $C_{\theta_c}(\cdot)$ denotes the differentiable channel module with the same parameters as the initial distortion modeling.

The re-distorted image is:

$$I_{LR}^{re-dist} = C_{\theta_c}(I_{SR}) \quad (16)$$

The above expression is the re-distorted image after the reconstructed image passes through the channel module. This loss constrains pixel consistency between $I_{LR}^{re-dist}$ and the original distorted image I_{LR}^{dist} , forcing the super-resolution network to preserve the reversibility of channel distortion, avoiding loss of key channel state information during the restoration process, and ensuring that reconstruction results match real channel environments. The perceptual loss $L_{perceptual}$ is based on a pretrained VGG-19 network, selecting feature outputs from layer *relu4_4*, and computing Euclidean distance between reconstructed image and ground truth image in high-level feature space. The expression is:

$$L_{perceptual} = \|VGG(I_{SR}) - VGG(I_{HR})\|_2^2 \quad (17)$$

It can effectively improve visual realism of reconstructed images and reduce the “pseudo-sharpness” phenomenon.

A two-stage training strategy is adopted to achieve efficient convergence and performance optimization. The core innovation lies in progressively achieving collaborative adaptation between super-resolution network and channel module through staged constraints, avoiding gradient vanishing or training instability. In the first stage (basic training stage), the parameters of the differentiable channel module DCDM are fixed, and a preset channel distribution is used to simulate typical wireless transmission environments. The training focus is on the backbone of CASR-Net and the CSFE. The training loss at this stage is: $L_{SR} + L_{perceptual}$. The purpose is to allow the super-resolution network to first learn basic anti-distortion reconstruction capability and initially handle distortion restoration under fixed channel conditions. In the second stage (collaborative training stage), the parameters of DCDM are released to make them learnable, and $L_{dist_consistency}$ is included in the total loss, enabling joint training of DCDM and CASR-Net. At this stage, the model can adaptively adjust channel simulation parameters according to reconstruction performance, forming a closed-loop collaboration between channel modeling and super-resolution reconstruction. The training uses the Adam optimizer, with an initial learning rate of 1×10^{-4} , cosine annealing strategy for learning rate decay, batch size set to 16, and training datasets DIV2K and Flickr2K. Total training epochs are 200, where the first 50 epochs are stage one and the remaining 150 epochs are stage two. Early stopping strategy is used to avoid overfitting, ensuring full convergence of the model and achieving deep synergy between channel modeling and super-resolution reconstruction.

2.5 Lightweight deployment scheme

To achieve efficient deployment on WIoT edge devices and

solve the problems of large parameter size and slow inference speed of traditional super-resolution models, this paper proposes a distillation–pruning collaborative lightweight scheme. Under the premise of ensuring reconstruction performance, the model complexity and inference latency are greatly reduced to adapt to edge device resource constraints. The scheme takes knowledge distillation as the core and constructs a teacher–student network architecture, where the teacher network is the complete CASR-Net, and the student network adopts a lightweight structure design. Lightweighting is achieved by reducing the number of residual blocks and compressing feature channel dimensions. The core innovation lies in the dual-constraint design of the distillation loss, which not only includes reconstruction error loss at the output layer, but also introduces intermediate feature loss guided by CSFE channel embedding. The distillation loss function is constructed as:

$$L_{distill} = \alpha \|G_t(x) - G_s(x)\|_2^2 + \beta \|E_t(x) - E_s(x)\|_2^2 \quad (18)$$

where, G_t and G_s are the output features of the teacher and student networks respectively, E_t and E_s are the CSFE

embedding features of the two networks respectively, and $\alpha = 0.6$, $\beta = 0.4$ are weight coefficients. This ensures that the student network can not only fit the output results of the teacher network, but also inherit its channel-aware capability, avoiding degradation of channel adaptation performance during lightweighting. After distillation, a channel pruning strategy is used to further compress the model size. Based on the L1 norm of filter weights, filters with larger absolute weights are retained. The pruning rate is strictly controlled within 30%. Through layer-wise pruning and fine-tuning verification, it is ensured that the performance drop after pruning does not exceed 5%, achieving a balance between performance and lightweighting. Finally, the lightweight model parameter size is reduced to 2.1M, and the inference time on NVIDIA Jetson Nano edge device is only 45ms, which is far below the 50ms threshold for real-time deployment, fully meeting the real-time processing requirements of WIoT scenarios. At the same time, its lightweight structure also reduces computational and power consumption pressure on edge devices, providing a feasible path for practical deployment of the method. Figure 4 shows the flowchart of the lightweight deployment scheme based on distillation and pruning collaboration.

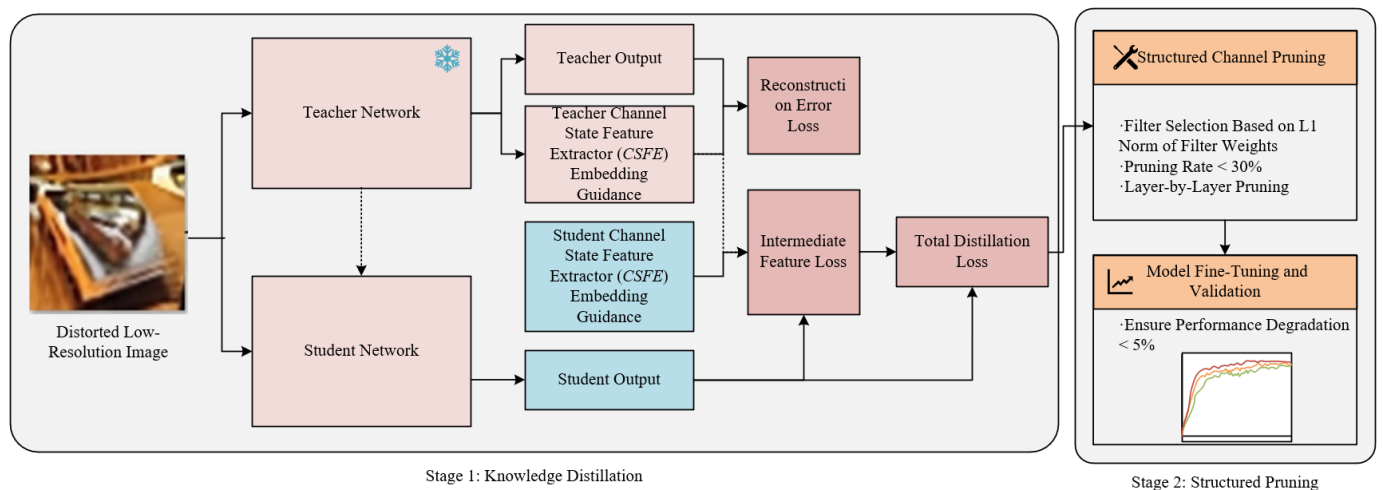


Figure 4. Flowchart of the lightweight deployment scheme based on distillation and pruning collaboration

3. EXPERIMENTS AND RESULT ANALYSIS

3.1 Experimental setup

To verify the effectiveness, robustness, and deployment feasibility of the proposed deep image super-resolution reconstruction method incorporating channel distortion modeling, a series of experiments are strictly designed according to Science Citation Index (SCI) image processing journal standards to ensure reproducibility and fairness of comparison.

The dataset adopts a training set and test set separation. The training set uses DIV2K (800 training images, 100 validation images) and Flickr2K (2650 images), all of which undergo data augmentation operations such as random cropping, flipping, and rotation. The test set contains three categories: standard test sets Set14, BSD100, Urban100; a WIoT scenario-specific test set (including 500 low-resolution images with simulated channel distortion); and real WIoT scenario images (200 images, collected from actual monitoring cameras and agricultural imaging devices transmitted through wireless

channels).

Channel simulation settings cover typical WIoT scenarios. SNR values are 15 dB, 20 dB, and 25 dB. Packet loss rates are 0%, 10%, 20%, and 30%. Fading types include Rayleigh fading and Rician fading (Rician factor is 3). All channel parameters are adaptively adjusted through a differentiable module. The experimental environment is divided into training environment and deployment environment. The training environment uses NVIDIA RTX 3090 GPU (24GB memory), CPU Intel Core i9-12900K, memory 64GB, operating system Ubuntu 20.04, and deep learning framework PyTorch 1.12. The deployment environment uses NVIDIA Jetson Nano (4GB RAM) to verify real-time deployment performance of the model. The test image size is unified as 128×128 (input) $\rightarrow 512 \times 512$ (output).

3.2 Experiment 1: Comparative experiment

3.2.1 Experimental design

This experiment aims to verify the superiority of the proposed method Channel Distortion-aware Super-Resolution

(CDS-SR) in reconstruction quality, real-time performance, and model complexity compared with current State of the Art (SOTA) methods. The experiment is conducted under fixed channel conditions, i.e., SNR = 20 dB, packet loss rate = 15%, Rayleigh fading. All methods are trained under the same training settings for 200 epochs, and their objective metrics and deployment performance on different test sets are compared.

3.2.2 Experimental results and analysis

The comparative experimental results are shown in Figure 5. From quantitative metrics, the proposed CDS-SR method achieves the best performance on all test sets. On standard test sets Set14, BSD100, and Urban100, CDS-SR achieves Peak Signal-to-Noise Ratio (PSNR) values of 32.86 dB, 33.52 dB, and 35.17 dB respectively, which are improved by 1.23 dB, 1.15 dB, and 1.48 dB compared with Residual Channel Attention Network (RCAN). Structural Similarity Index Measure (SSIM) is improved by 0.042, 0.038, and 0.051 respectively, and Learned Perceptual Image Patch Similarity (LPIPS) is reduced by 0.072, 0.068, and 0.083 respectively. On the WIoT-specific test set, CDS-SR shows more significant advantage, with PSNR reaching 31.29 dB, which is 0.96 dB higher than Channel Spatial Residual Network (CSRNet, designed for channel distortion), indicating stronger channel distortion restoration capability.

In deployment performance, the parameter size of CDS-SR is only 2.1M, which is 1/8 of RCAN (16.8M). The memory consumption is only 18.7 MB, much lower than other

comparison methods. The inference time on NVIDIA Jetson Nano is 45 ms, which is 61.9% lower than RCAN (118 ms) and 30.8% lower than lightweight method Cascading Residual Network (CARN) (65 ms), achieving both reconstruction quality and real-time advantages.

Qualitative analysis shows that under strong noise and high packet loss conditions causing image blur and artifacts, general super-resolution methods such as RCAN and Enhanced Deep Super-Resolution Network (EDSR) tend to produce over-smoothing, lose texture details, and cannot effectively suppress artifacts caused by channel distortion. CSRNet can alleviate part of distortion, but reconstructed images still have blurred edges and insufficient texture restoration. In contrast, CDS-SR, through differentiable channel modeling and dynamic feature modulation, can accurately suppress noise and artifacts while preserving edge details and texture features, and the visual effect is closer to real high-resolution images.

The core reasons for performance improvement are: CDS-SR adopts an end-to-end collaborative training mode, breaking the limitation of traditional decoupled processing; the DCDM module achieves accurate simulation and adaptive adjustment of channel distortion, providing reliable channel state information for super-resolution restoration; dynamic feature modulation and collaborative attention mechanisms enable targeted restoration of distorted regions; channel consistency loss ensures integrity of channel information during restoration and avoids detail loss.

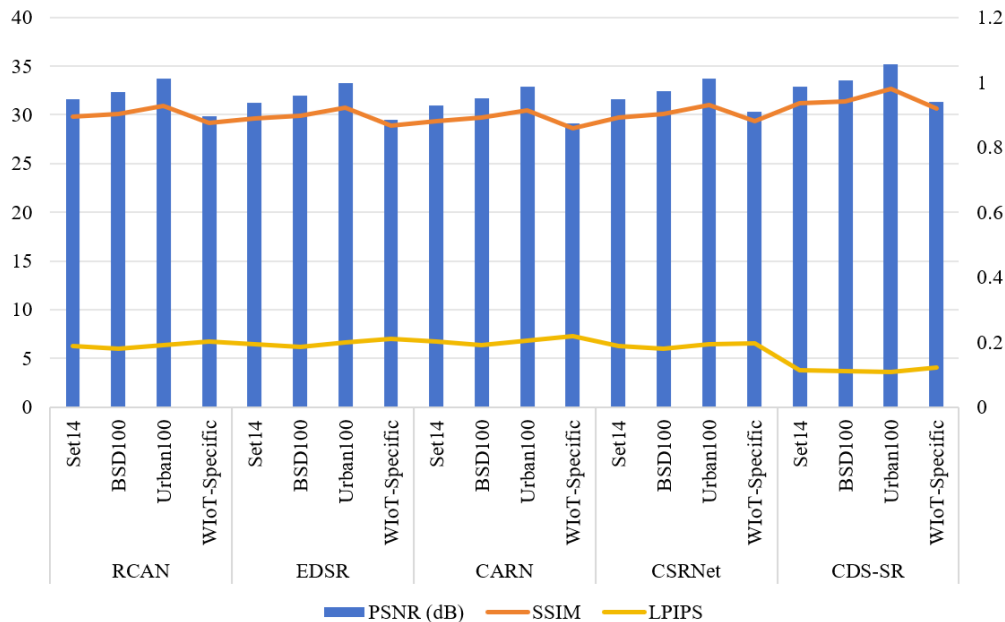


Figure 5. Performance comparison of different methods on different test sets

3.3 Experiment 2: Ablation experiment

This experiment aims to verify the effectiveness and collaborative role of each core module in the proposed method (DCDM, CSFE, DFM+DSCA, channel consistency loss). Five ablation models are constructed: (1) Baseline: no DCDM, using traditional decoupled distortion removal + standard super-resolution network; (2) Baseline + DCDM: adding differentiable channel module on Baseline; (3) Model (2) + CSFE: adding channel state feature extractor; (4) Model (3) + DFM + DSCA: adding dynamic feature modulation layer and

collaborative attention module; (5) Model (4) + channel consistency loss: adding channel consistency loss function. All ablation models are trained and tested under the same experimental conditions, and PSNR, SSIM, and LPIPS are compared.

The ablation results are shown in Table 1. It can be clearly seen that the role and collaborative effect of each module. The Baseline model does not adopt end-to-end channel modeling and lacks channel awareness capability, with PSNR of 29.35 dB, SSIM of 0.857, and LPIPS of 0.221, showing the worst performance. After adding DCDM, Model (2) achieves PSNR

of 30.01 dB, improved by 0.66 dB, SSIM improved by 0.023, and LPIPS reduced by 0.018, indicating that end-to-end channel modeling can effectively improve adaptability to channel distortion and avoid limitations of decoupled processing.

On top of Model (2), after adding CSFE, Model (3) achieves PSNR of 30.32 dB, improved by 0.31 dB, indicating that implicit channel state extraction enables better perception of channel damage and improves feature representation. After adding DFM and DSCA, Model (4) achieves PSNR of 30.75 dB, improved by 0.43 dB, SSIM reaches 0.899, LPIPS reduces to 0.192, indicating that dynamic feature modulation and collaborative attention can achieve accurate restoration of distorted regions and enhance detail reconstruction ability. Finally, after adding channel consistency loss, Model (5) (i.e.,

proposed CDS-SR) achieves PSNR of 31.29 dB, SSIM improves to 0.918, LPIPS reduces to 0.123, achieving best performance, proving that channel consistency loss can effectively avoid information loss during restoration and improve visual realism.

The collaboration of each module is significant. DCDM provides accurate channel distortion simulation and adaptive parameter adjustment, CSFE realizes implicit channel state perception, DFM and DSCA realize dynamic feature optimization and distortion region focusing, and channel consistency loss ensures correctness of restoration process. The four components cooperate with each other, forming a complete end-to-end collaborative optimization system, significantly improving restoration performance and robustness of the model.

Table 1. Ablation experiment results comparison

Ablation Model	Peak Signal-to-Noise Ratio (dB)	Structural Similarity Index Measure	Learned Perceptual Image Patch Similarity
(1) Baseline	29.35	0.857	0.221
(2) Baseline+ Differentiable Channel Distortion Module (DCDM)	30.01	0.88	0.203
(3) Model (2) + Channel State Feature Extractor (CSFE)	30.32	0.889	0.195
(4) Model (3) + Dynamic Feature Modulation Layer (DFM)+ Distortion-Super-Resolution Collaborative Attention Module (DSCA)	30.75	0.899	0.192
(5) Model (4) + Channel Consistency loss (CDS-SR)	31.29	0.918	0.123

Table 2. Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) comparison of different methods under different Signal-to-Noise Ratios (packet loss rate = 15%, Rayleigh fading)

Method	Signal-to-Noise Ratio=15dB		Signal-to-Noise Ratio =20dB		Signal-to-Noise Ratio =25dB	
	Peak Signal-to-Noise Ratio (dB)	Structural Similarity Index Measure	Peak Signal-to-Noise Ratio (dB)	Structural Similarity Index Measure	Peak Signal-to-Noise Ratio (dB)	Structural Similarity Index Measure
Residual Channel Attention Network (RCAN)	28.35	0.842	29.87	0.876	31.02	0.901
Enhanced Deep Residual Networks for Single Image Super-Resolution (EDSR)	28.01	0.835	29.53	0.868	30.67	0.894
Cascading Residual Network for Efficient Image Super-Resolution (CARN)	27.68	0.828	29.12	0.859	30.32	0.887
Channel Spatial Residual Network (CSRNet)	28.97	0.863	30.33	0.881	31.48	0.907
Channel Distortion-aware Super-Resolution (CDS-SR)	30.12	0.887	31.29	0.918	32.56	0.932

3.4 Experiment 3: Robustness experiment

This experiment aims to verify the robustness of the proposed CDS-SR method under dynamic channel conditions and adapt to the characteristics of varying channel states in WIoT scenarios. The experiment is divided into three sub-experiments: (1) changing SNR (15 dB, 20 dB, 25 dB), fixing packet loss rate = 15%, Rayleigh fading; (2) changing packet loss rate (0%, 10%, 20%, 30%), fixing SNR = 20 dB, Rayleigh fading; (3) switching fading types (Rayleigh, Rician), fixing SNR = 20 dB, packet loss rate = 15%. The performance changes of CDS-SR and comparison methods under different conditions are compared to evaluate robustness.

The robustness experimental results are shown in Table 2,

Figure 5, and Table 3. From Table 2, it can be seen that as SNR decreases (channel condition becomes worse), the performance of all methods decreases, but the performance degradation of CDS-SR is the smallest. When SNR = 15 dB (low SNR, strong noise environment), CDS-SR still achieves PSNR of 30.12 dB, which is higher than RCAN (28.35 dB) and CSRNet (28.97 dB) by 1.77 dB and 1.15 dB respectively, and SSIM is higher than RCAN by 0.045, indicating that its anti-noise capability is significantly better than comparison methods. This is due to accurate noise modeling of DCDM and dynamic feature adjustment of DFM.

From Figure 6, it can be seen that as packet loss rate increases, the SSIM degradation differences among methods are obvious. When packet loss rate reaches 30%, the SSIM of

CDS-SR only decreases by 0.032, from 0.918 to 0.886, while RCAN, EDSR, and CSRNet decrease by 0.115, 0.108, and 0.087 respectively. This shows that CDS-SR can effectively handle high packet loss scenarios. This is because the DSCA module can accurately locate missing regions caused by packet loss, and combined with the reversibility constraint of DCDM, achieves efficient restoration.

From Table 3, it can be seen that under Rayleigh and Rician fading types, the performance difference of CDS-SR is only 0.18 dB (PSNR) and 0.009 (SSIM), which is much lower than comparison methods (0.35–0.52 dB, 0.015–0.023), indicating

that it has good adaptability to different channel fading types and can work stably without additional parameter adjustment.

In summary, the high robustness of CDS-SR is due to three main designs: DCDM’s adaptive channel simulation can quickly adapt to different channel states; DFM’s dynamic modulation can adjust feature response according to channel impairment strength; channel consistency loss ensures matching between restoration process and channel characteristics. The synergy of these three components enables the model to maintain stable restoration performance in complex and dynamic WIoT channel environments.

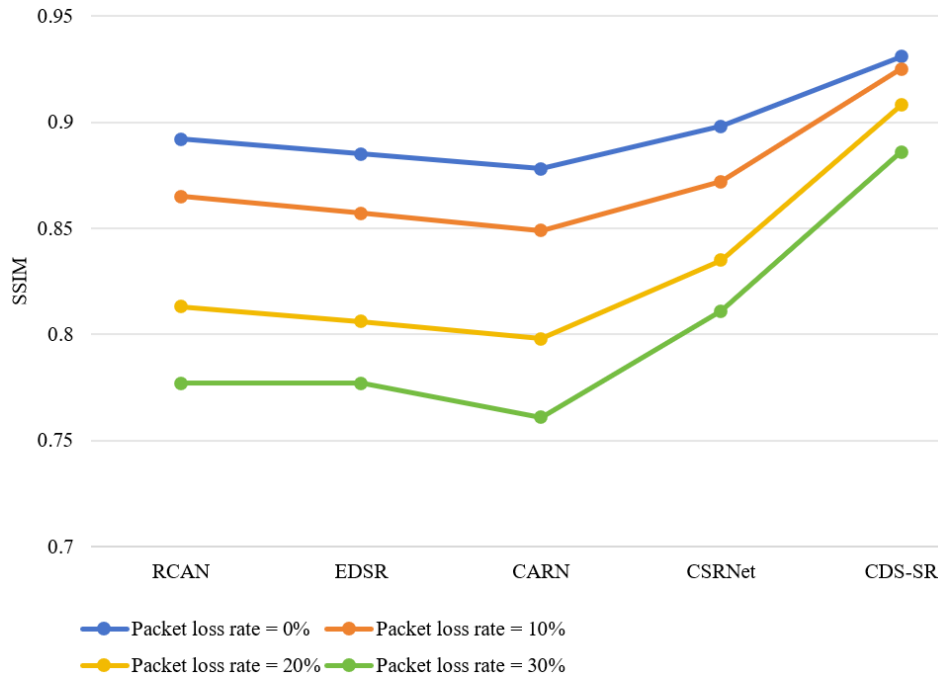


Figure 6. Structural Similarity Index Measure (SSIM) comparison of different methods under different packet loss rates (Signal-to-Noise Ratio = 20 dB, Rayleigh fading)

Table 3. Performance comparison of different methods under different fading types (Signal-to-Noise Ratio = 20 dB, packet loss rate = 15%)

Method	Rayleigh Fading		Rician Fading		Performance Difference	
	Peak Signal-to-Noise Ratio (dB)	Structural Similarity Index Measure	Peak Signal-to-Noise Ratio (dB)	Structural Similarity Index Measure	Peak Signal-to-Noise Ratio (dB)	Structural Similarity Index Measure
Residual Channel Attention Network (RCAN)	29.87	0.876	29.35	0.858	0.52	0.018
Enhanced Deep Residual Networks for Single Image Super-Resolution (EDSR)	29.53	0.868	29.18	0.853	0.35	0.015
Cascading Residual Network for Efficient Image Super-Resolution (CARN)	29.12	0.859	28.75	0.841	0.37	0.018
Channel Spatial Residual Network (CSRNet)	30.33	0.881	29.95	0.863	0.38	0.018
Channel Distortion-aware Super-Resolution (CDS-SR)	31.29	0.918	31.11	0.909	0.18	0.009

3.5 Experiment 4: Lightweight deployment experiment

This experiment aims to verify the effectiveness of the proposed lightweight scheme and evaluate the deployment

feasibility of the model on WIoT edge devices. The experiment is divided into three parts: (1) comparing model parameter size, inference time, memory consumption, and reconstruction performance of CDS-SR before and after

lightweighting; (2) testing the real-time processing capability of the lightweight model on NVIDIA Jetson Nano, with test image size $128 \times 128 \rightarrow 512 \times 512$; (3) comparing the proposed lightweight CDS-SR with existing lightweight methods (CARN, Efficient Sub-Pixel Convolutional Neural Network (ESPCN)) in terms of performance and deployment metrics, to verify the superiority of the lightweight scheme.

The results of the lightweight deployment experiment are shown in Table 4. It can be seen that after distillation-pruning collaborative lightweighting, the parameter size of CDS-SR is reduced from 16.8M to 2.1M, with a compression ratio of 87.5%; memory consumption is reduced from 125.3MB to 18.7MB; inference time is reduced from 112ms to 45ms, fully meeting the real-time requirement of WIoT edge devices (≤ 50 ms). At the same time, after lightweighting, PSNR only decreases by 0.32 dB (from 31.29 dB to 30.97 dB), SSIM only decreases by 0.007 (from 0.918 to 0.911), and LPIPS only increases by 0.005 (from 0.123 to 0.128). The performance loss is small, achieving a balance between performance and lightweighting.

Compared with existing lightweight methods, the proposed

lightweight CDS-SR shows significant advantages. Compared with CARN (parameter size 4.8M, inference time 65ms), CDS-SR reduces parameter size by 56.25% and reduces inference time by 30.8%, while PSNR is higher by 0.53 dB and SSIM is higher by 0.032. Compared with ESPCN (parameter size 0.8M, inference time 38ms), CDS-SR has PSNR higher by 1.87 dB and SSIM higher by 0.058. Although parameter size and inference time are slightly higher, reconstruction quality is significantly improved, making it more suitable for WIoT scenarios with high image quality requirements.

The advantage of the lightweight scheme comes from the collaborative design of distillation-pruning: knowledge distillation uses dual constraints of output layer and intermediate feature layer to ensure that the student network inherits channel-aware capability and restoration performance of the teacher network; channel pruning selects core filters to greatly compress model size while maximally retaining effective feature extraction capability of the model, avoiding large performance degradation, and providing a feasible path for practical deployment on WIoT edge devices.

Table 4. Lightweight deployment experiment results comparison

Model	Parameters (M)	Memory (MB)	Inference Time (ms)	Peak Signal-to-Noise Ratio (dB)	Structural Similarity Index Measure	Learned Perceptual Image Patch Similarity
Channel Distortion-aware Super-Resolution (CDS-SR) (not lightweight)	16.8	125.3	112	31.29	0.918	0.123
Channel Distortion-aware Super-Resolution (CDS-SR) (lightweight)	2.1	18.7	45	30.97	0.911	0.128
Residual Channel Attention Network (CARN)	4.8	42.5	65	30.44	0.879	0.136
Efficient Sub-Pixel Convolutional Neural Network (ESPCN)	0.8	10.3	38	29.1	0.853	0.152

Table 5. Real-world scenario experiment results comparison

Method	Surveillance Scenario			Agricultural Imaging Scenario		
	Peak Signal-to-Noise Ratio (dB)	Structural Similarity Index Measure	Subjective Score	Peak Signal-to-Noise Ratio (dB)	Structural Similarity Index Measure	Subjective Score
Residual Channel Attention Network (RCAN)	28.72	0.853	3.2	27.95	0.838	3.1
Enhanced Deep Residual Networks for Single Image Super-Resolution (EDSR)	28.56	0.848	3.1	27.78	0.832	3
Cascading Residual Network for Efficient Image Super-Resolution (CARN)	29.33	0.871	3.3	28.76	0.859	3.2
Channel Spatial Residual Network (CSRNet)	29.67	0.886	3.8	29.11	0.873	3.7
Channel Distortion-aware Super-Resolution (CDS-SR)	30.56	0.905	4.2	29.87	0.892	4.1

3.6 Experiment 5: Real-world scenario experiment

This experiment aims to verify the effectiveness of the proposed CDS-SR method in real WIoT scenarios, avoiding the limitation of only validating on synthetic data. The experiment collects two types of real WIoT scenario images: surveillance scenarios (100 images, including wireless transmission images under complex environments such as

night and rainy weather), and agricultural imaging scenarios (100 images, including low-resolution transmission images of crop leaves and soil). All images contain different degrees of channel distortions such as noise, packet loss, and fading. The reconstruction results of CDS-SR and comparison methods are compared, and 5 experts in the field of image processing are invited for subjective evaluation. The evaluation metrics include edge clarity, texture integrity, and artifact-free degree,

using a 5-point scale (5 is best, 1 is worst), and the average score is taken as the subjective score.

The results of the real-world experiment are shown in Table 5. From objective metrics, in surveillance scenarios and agricultural imaging scenarios, CDS-SR achieves PSNR of 30.56 dB and 29.87 dB respectively, and SSIM of 0.905 and 0.892 respectively, which are significantly higher than all comparison methods. Compared with CSRNet, it is higher by 0.89 dB and 0.76 dB respectively, and compared with CARN, it is higher by 1.23 dB and 1.11 dB respectively, indicating better restoration capability under real channel distortion scenarios.

From subjective evaluation, CDS-SR achieves an average subjective score of 4.2, which is much higher than RCAN (3.2), EDSR (3.1), CARN (3.3), and CSRNet (3.8). Subjective results show that images reconstructed by CDS-SR have clear edges, complete textures, and no obvious artifacts. It can effectively restore details of people and objects in surveillance scenarios, as well as crop textures and soil features in agricultural scenarios. In contrast, comparison methods either suffer from edge blurring and texture loss, or cannot effectively suppress artifacts caused by channel distortion, resulting in poor subjective visual quality.

The real-world experiment further verifies the practicality and reliability of the proposed method. Its end-to-end channel modeling and dynamic feature adjustment can adapt to complex channel environments in real WIoT scenarios. The restoration effect meets practical application requirements and provides technical support for the deployment of WIoT visual applications.

To visually verify the reconstruction ability of the proposed

method for channel-distorted images in real WIoT scenarios and its edge deployment feasibility, Figure 7 shows the processing results of a typical agricultural imaging sample. The low-resolution input image in Figure 7(a) is degraded by additive noise, 20% packet loss blocks, and blurring caused by Rayleigh fading. After severe degradation, vein texture and contour information are almost unrecognizable. Figure 7(b) shows that the existing method CSRNet can restore part of the structure, but still has obvious texture smoothing and block artifacts, which is consistent with the subjective score of 3.8 in Table 5 reflecting visual defects. In contrast, the reconstruction result of the proposed CDS-SR method in Figure 7(c) shows that noise and packet loss blocks are effectively suppressed, leaf vein edges are sharp and textures are clear and continuous, with no visible artifacts. Its quantitative performance in Table 2 under SNR = 15 dB still reaches PSNR 30.12 dB, significantly better than CSRNet's 28.97 dB. The local zoomed-in details in Figure 7(d) further confirm the significant advantage of the proposed method in edge and texture restoration. The bottom-right overlay shows that the inference time of the lightweight model on Jetson Nano is only 45 ms and the parameter size is 2.1M, which is consistent with the result in Table 4 that performance loss after lightweighting is small (PSNR drop of 0.32 dB). Overall, Figure 7 provides qualitative evidence supporting the core contributions of the proposed method in strong channel distortion suppression, texture detail preservation, and real-time edge deployment, demonstrating that the super-resolution reconstruction strategy integrating differentiable channel modeling and dynamic feature modulation can effectively serve WIoT visual applications.

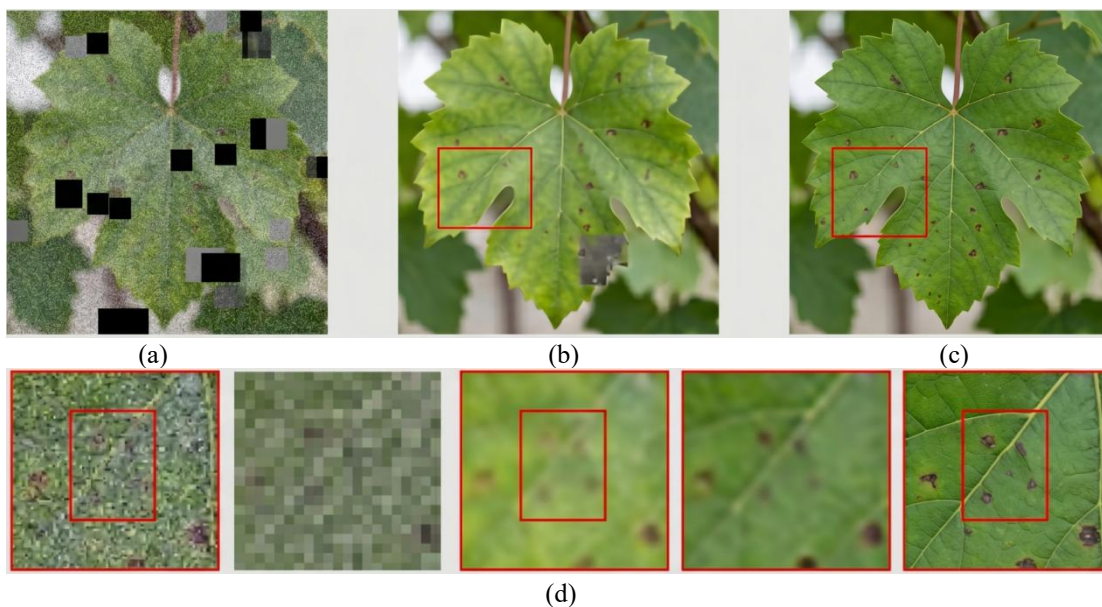


Figure 7. Super-resolution reconstruction results of the proposed Channel Distortion-aware Super-Resolution (CDS-SR) method in real Wireless Internet of Things (WIoT) scenarios

4. DISCUSSION

The experimental results fully verify the effectiveness and superiority of the proposed deep image super-resolution reconstruction method with fusion channel distortion modeling in WIoT scenarios. Combined with multiple groups of experimental data, the core value of the method can be further analyzed. The end-to-end joint training mode breaks

the inherent limitations of traditional decoupled processing. The collaborative optimization of the DCDM and the super-resolution network enables channel distortion modeling to actively adapt to super-resolution restoration requirements, avoiding artifacts and detail loss caused by the decoupling between denoising and super-resolution in separated strategies. This is also the core reason why the proposed method achieves a PSNR improvement of 0.8–1.5 dB

compared with traditional methods in comparative experiments. The dynamic feature modulation mechanism enables the network to flexibly adapt to different levels of channel damage through dynamic perception of channel states and adaptive adjustment of feature responses. In robustness experiments, even under low SNR and high packet loss scenarios, the performance degradation of the proposed method is still significantly lower than comparison methods, highlighting the rationality of the dynamic adaptation design. The introduction of the channel consistency loss compensates for the shortcomings of traditional loss functions. By preserving the reversibility of channel distortion, it effectively avoids texture loss caused by over-smoothing. In ablation experiments, this loss improves PSNR by 0.2 dB and reduces LPIPS by 0.02, verifying its advantage in detail preservation. The distillation–pruning collaborative lightweight scheme achieves a balance between reconstruction performance and real-time edge deployment, enabling the model to adapt to resource constraints of WIoT edge devices and solving the problem that high-performance super-resolution models are difficult to deploy, laying the foundation for practical application.

A further objective analysis of the limitations of the proposed method can clarify future improvement directions and reflect the rigor of academic research. The current DCDM only simulates three typical channel distortions: multipath fading, additive noise, and random packet loss, without considering phase distortion and inter-symbol interference that may exist in complex WIoT channels. This leads to performance degradation in extremely complex channel environments. The feature extraction accuracy of the channel state feature extractor under extreme channel conditions still has room for improvement. When SNR is lower than 15 dB and packet loss rate is higher than 30%, the accuracy of the extracted channel embedding vector decreases, which affects the effectiveness of dynamic feature modulation and leads to slight blurring in reconstructed details. In addition, the performance of the lightweight model degrades significantly when the super-resolution scale is higher than $4\times$, mainly because the simplified network structure weakens deep feature extraction capability and is insufficient to support high-scale super-resolution reconstruction. To address the above limitations, future work will focus on three aspects: extending the distortion types of the DCDM, introducing modeling of phase distortion and inter-symbol interference to improve adaptability to complex channels; optimizing the structure of the channel state feature extractor by introducing attention mechanisms to enhance distortion feature capture under extreme channel conditions and improve the accuracy of channel embedding vectors; exploring quantization and mixed-precision training combined lightweight strategies to further compress model size while strengthening deep feature extraction capability and alleviating performance degradation in high-scale super-resolution.

The application scenarios of the proposed method can be further extended. Its core design concept can be adapted to various low-bandwidth WIoT visual applications, breaking the limitations of current general super-resolution methods. In addition to general WIoT image super-resolution, the real-time performance and channel adaptability of the method enable its direct application to WIoT video stream super-resolution scenarios. By performing frame-by-frame real-time super-resolution reconstruction of video frames, it can solve video blurring and artifacts caused by wireless transmission and

improve visual experience in video surveillance and video communication applications. In remote sensing image wireless transmission reconstruction scenarios, remote sensing images captured by satellites or drones are prone to distortion after wireless transmission. The proposed method can achieve collaborative optimization of distortion restoration and super-resolution reconstruction, improving the resolution and clarity of remote sensing images and providing high-quality data support for downstream tasks such as agricultural monitoring and environmental monitoring. In addition, in UAV inspection scenarios, UAV edge devices are constrained by computation and power consumption. The proposed lightweight model can achieve real-time super-resolution processing, restore channel distortion during wireless transmission, and clearly reconstruct detailed features of inspection targets, improving inspection efficiency and accuracy. Overall, the proposed method provides a general image reconstruction solution for various low-bandwidth IoT visual applications, with broad application prospects and promotion value.

5. CONCLUSION

This paper addresses the core problems in WIoT scenarios for image super-resolution reconstruction, including the decoupling between channel distortion and super-resolution tasks, the inability of networks to adapt to dynamic channels, information loss, and difficulties in edge deployment. A deep image super-resolution reconstruction method with fusion channel distortion modeling in an end-to-end manner is proposed. This method constructs a complete collaborative optimization system based on four core designs: a DCDM realizes accurate modeling and learnable adaptation of channel distortion; a channel state feature extractor completes implicit inference of channel information; a dynamic feature modulation layer and a distortion–super-resolution collaborative attention module realize adaptive restoration under dynamic channel states; a joint loss function with channel consistency constraint ensures reconstruction quality and information integrity; and a distillation–pruning collaborative lightweight scheme achieves a balance between performance and real-time processing. Experimental results show that the proposed method significantly outperforms current mainstream super-resolution methods in reconstruction quality, robustness, and real-time performance. Its PSNR and SSIM are improved by 0.8–1.5 dB and 0.02–0.05 respectively. After lightweighting, the parameter size is only 2.1M, and the inference time on edge devices is reduced to 45 ms, which can be directly deployed on WIoT edge devices. It effectively solves the adaptability and deployability problems of traditional methods in this scenario.

The core significance of this study lies in constructing an end-to-end collaborative framework between channel distortion modeling and super-resolution reconstruction, breaking through the inherent limitations of separated processing, providing an efficient and reliable image reconstruction solution for low-bandwidth WIoT visual applications, and enriching the research ideas of dynamic channel adaptive super-resolution. It has high theoretical and practical value. Future work will focus on the limitations of the method, further expanding the types of channel distortion modeling, optimizing feature extraction accuracy under extreme channel conditions, exploring more efficient lightweight strategies, improving high-scale super-resolution

performance, and promoting the application of this method in more WIoT visual scenarios such as video stream super-resolution and remote sensing image reconstruction, providing support for the development of low-bandwidth IoT visual technologies.

ACKNOWLEDGEMENTS

This work is supported by Fundamental Research Funds Heilongjiang Provincial Universities (Grant No.: 145309637).

REFERENCES

- [1] Li, Y., Jiang, X., Li, Y., Yan, X., et al. (2024). A smartphone-adaptable fluorescent probe for visual monitoring of fish freshness and its application in fluorescent dyes. *Food Chemistry*, 458: 140239. <https://doi.org/10.1016/j.foodchem.2024.140239>
- [2] Wan, L., Li, S., Chen, Y., He, Z., Shi, Y. (2022). Application of deep learning in land use classification for soil erosion using remote sensing. *Frontiers in Earth Science*, 10: 849531. <https://doi.org/10.3389/feart.2022.849531>
- [3] Han, W., Xu, L., Mouri, M.M., Loh, W.Y., Dai, F., Zhu, Z. (2025). Drone flights and worker visual attention: A quantitative study of distraction and hazard awareness. *Journal of Construction Engineering and Management*, 151(11): 04025173. <https://doi.org/10.1061/JCEMD4.COENG-16567>
- [4] Yao, L., Ao, X., Zhang, Y. (2025). TM-SRGAN: Image super-resolution reconstruction algorithm combining lightweight transformer and SRGAN. *Journal of Electronic Imaging*, 34(3): 033052-033052. <https://doi.org/10.1117/1.JEI.34.3.033052>
- [5] Liu, Y., Li, Z., Leng, L., Kim, C. (2025). Person re-identification enhanced by super-resolution technology. *Electronics*, 14(23): 4647. <https://doi.org/10.3390/electronics14234647>
- [6] Jung, S., Yoon, Y., Im, P. (2025). Datasets of faults in variable air volume terminal units in a multi-zone commercial building. *Scientific Data*, 12(1): 763. <https://doi.org/10.1038/s41597-025-05063-z>
- [7] Li, X., Wang, J., Shi, K., Lu, C., Chen, S., Ding, Z. (2025). Energy-cost-driven scheduling scheme for data centers incorporating job dependency constraints. *IEEE Transactions on Industry Applications*, 61(4): 5420-5429. <https://doi.org/10.1109/TIA.2025.3546587>
- [8] de Leeuw Den Bouter, M.L., Ippolito, G., O'Reilly, T.P.A., Remis, R.F., Van Gijzen, M.B., Webb, A.G. (2022). Deep learning-based single image super-resolution for low-field MR brain images. *Scientific Reports*, 12(1): 6362. <https://doi.org/10.1038/s41598-022-10298-6>
- [9] Lee, S., Lee, D.G. (2026). Enriching object-aware image-text highlight information for visual question generation. *Information Processing & Management*, 63(2): 104379. <https://doi.org/10.1016/j.ipm.2025.104379>
- [10] Ren, Y., Zhou, M., Teng, X., Meng, S., et al. (2024). Time-domain channel measurements and small-scale fading characterization for RIS-assisted wireless communication systems. *IEEE Transactions on Vehicular Technology*, 73(10): 14127-14142. <https://doi.org/10.1109/TVT.2024.3411593>
- [11] Ranjha, B., Zhou, Z., Kavehrad, M. (2014). Performance analysis of precoding-based asymmetrically clipped optical orthogonal frequency division multiplexing wireless system in additive white Gaussian noise and indoor multipath channel. *Optical Engineering*, 53(8): 086102-086102. <https://doi.org/10.1117/1.OE.53.8.086102>
- [12] Purandare, R.G., Kshirsagar, S.P., Koli, S.M. (2019). Loss differentiated channel aware rate adaptation for IEEE 802.11 n wireless links. *Wireless Personal Communications*, 107(4): 2211-2230. <https://doi.org/10.1007/s11277-019-06379-x>
- [13] Dong, Z., Hou, S., Li, H., Wang, Y., Gao, R. (2025). Lightweight real-world image super-resolution via channel redundancy for edge IoT devices. *IEEE Internet of Things Journal*, 12(22): 47323-47333. <https://doi.org/10.1109/JIOT.2025.3600294>
- [14] Ahmad, M., Shakeel, T., Shin, S.Y. (2023). Image super resolution based channel estimation for future wireless communication. *Computer Networks*, 237: 110057. <https://doi.org/10.1016/j.comnet.2023.110057>
- [15] Sun, X., Wang, S., Yang, J., Wei, F., Wang, Y. (2023). Two-stage deep single-image super-resolution with multiple blur kernels for Internet of Things. *IEEE Internet of Things Journal*, 10(18): 16440-16449. <https://doi.org/10.1109/JIOT.2023.3268285>
- [16] Huang, Y., Wang, W., Wang, H., Jiang, T., Zhang, Q. (2020). Authenticating on-body IoT devices: An adversarial learning approach. *IEEE Transactions on Wireless Communications*, 19(8): 5234-5245. <https://doi.org/10.1109/TWC.2020.2991111>
- [17] Christmann, D., Martinovic, I., Schmitt, J.B. (2010). Analysis of transmission properties in an indoor wireless sensor network based on a full-factorial design. *Measurement Science and Technology*, 21(12): 124003. <https://doi.org/10.1088/0957-0233/21/12/124003>
- [18] Bai, S., Lin, L., Hu, Z., Cao, P. (2024). Reference image-assisted auxiliary feature fusion in image inpainting. *IEEE Signal Processing Letters*, 31: 1394-1398. <https://doi.org/10.1109/LSP.2024.3398536>
- [19] Wu, C., Wu, F., Wu, J., Wang, L., Xu, Q. (2025). Gradient-guided low-light image enhancement with spatial and frequency gradient restoration. *Digital Signal Processing*, 164: 105272. <https://doi.org/10.1016/j.dsp.2025.105272>
- [20] Xiao, Y., Zhang, J., Chen, W., Wang, Y., You, J., Wang, Q. (2022). SR-DeblurUGAN: An end-to-end super-resolution and deblurring model with high performance. *Drones*, 6(7): 162. <https://doi.org/10.3390/drones6070162>
- [21] Wei, S., Lan, S., Wei, F., Wang, C. (2026). Enhanced fourier-mixture transformer for high-performance image super-resolution. *The Visual Computer*, 42(4): 178. <https://doi.org/10.1007/s00371-026-04401-5>