International Journal of Basic and Applied Sciences, 14 (4) (2025) 17-23

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/vv8ve405
Research paper

Comparative Analysis of Integrated Learning Based Methods for Image Super-Resolution

Kavya T. M. Research Scholar [1] *, Dr. Yogish Naik G. R. [2]

[1] Department of P.G Studies and Research in Computer Science, Kuvempu University, Shivamogga, Karnataka, India
[2] Associate Professor, Department of P.G Studies and Research in Computer Science, Kuvempu University, Shivamogga, Karnataka, India
**Corresponding author E-mail: kavyatm29@gmail.com*

Received: June 23, 2025, Accepted: July 28, 2025, Published: August 2, 2025

Abstract

Image super-resolution (SR) is a fundamental problem in computer vision aimed at reconstructing high-resolution (HR) images from low-resolution (LR) inputs. Various methods have been proposed to solve this problem, ranging from traditional signal processing techniques to modern deep learning-based approaches. Two prominent methods in SR are Sparse Representation (SR) and the ASDS_AR (Attention-based Super-Resolution with Discriminative Spatial Attention Residuals) technique. Sparse representation, typically used with dictionary learning and sparse coding, models images as sparse combinations of learned atoms from a dictionary, allowing for the recovery of fine details. On the other hand, the ASDS_AR method leverages deep learning, particularly attention mechanisms and residual learning, to focus on important spatial regions and enhance the image resolution. This paper presents a comparative study of these two approaches in terms of their performance, computational efficiency, and ability to preserve fine image details. Through qualitative and quantitative evaluation on several benchmark datasets, we analyze the strengths and weaknesses of each method in terms of image quality, PSNR, SSIM, and perceptual metrics.

Keywords: *Sparse Representation; ASDS_AR; K_SVD; SRGAN.*

1. Introduction

The task of image super-resolution involves reconstructing high-resolution (HR) images from their low-resolution (LR) counterparts, which has a wide range of applications in medical imaging, satellite imaging, security systems, and multimedia. Over the years, two prominent approaches have gained attention in the SR community: sparse representation and deep learning-based methods such as ASDS_AR.

- Sparse Representation (SR) is a mathematical approach that models an image as a sparse linear combination of elements from an over-complete dictionary. By learning the dictionary of atoms, sparse representation-based SR methods aim to recover fine details and high-frequency components of the image. Popular methods include the use of sparse coding, K-SVD, and matching pursuit algorithms.
- ASDS_AR (Attention-based Super-Resolution with Discriminative Spatial Attention Residuals) is a deep learning-based method that combines residual learning with attention mechanisms. ASDS_AR uses attention to focus on crucial spatial areas, dynamically guiding the network toward enhancing important regions of the image while preserving structural integrity.

This paper explores the performance of sparse representation and ASDS_AR in image super-resolution. The goal is to assess which method is more effective at reconstructing high-quality HR images and to identify the contexts where each method excels.

2. Background and literature review

2.1. Sparse representation for image super-resolution

Sparse representation methods have been widely used in image processing, particularly in SR. The key idea is to represent the high-resolution image as a sparse combination of dictionary elements that capture the image's underlying structure. The typical workflow of sparse-based SR involves:

1) Dictionary Learning: Constructing an over-complete dictionary where each atom represents an image patch or feature.
2) Sparse Coding: The image is then represented as a sparse linear combination of the learned dictionary elements.
3) Image Reconstruction: Using the sparse codes to reconstruct the high-resolution image through either direct optimization or iterative refinement.

Some notable methods include:

- K-SVD (K-means Singular Value Decomposition): A popular algorithm for dictionary learning.
- Matching Pursuit: A greedy algorithm for sparse coding that selects dictionary elements that best match the input image patch.
- LASSO-based Optimization: Regularized least squares optimization used for sparse coding.

Sparse representation has the advantage of being interpretable and capable of recovering fine image details. However, it suffers from high computational complexity due to the need for dictionary learning and sparse coding, making it less suitable for real-time applications.

2.2. ASDS_AR method

The ASDS_AR method is a deep learning-based approach that employs a residual learning framework combined with attention mechanisms for image super-resolution. The model architecture is designed to focus on learning the residual (difference) between the low-resolution and high-resolution images. This allows the network to recover fine details that might be lost in traditional SR methods.
Key components of ASDS_AR include:

- Residual Learning: The network learns the residual, which is the difference between the high-resolution and low-resolution images, enabling the model to focus on high-frequency components (details).
- Discriminative Spatial Attention: The model uses attention mechanisms to identify and focus on the most important regions of the image, such as edges and textures.
- End-to-End Deep Learning: The model is trained end-to-end, allowing it to learn from data without the need for pre-trained dictionaries or manual feature engineering.

ASDS_AR is highly flexible and can be adapted for different image types and applications. It is also computationally efficient compared to traditional sparse representation methods due to its ability to leverage modern hardware (GPUs) for parallel processing.

In the field of image super-resolution (SR), researchers have long sought to develop techniques that can enhance the quality of low-resolution (LR) images to generate high-resolution (HR) versions. Over the years, a wide range of methods have emerged, from traditional interpolation techniques to sophisticated learning-based deep neural networks (DNNs). Hybrid methods combine the advantages of multiple SR techniques, typically blending traditional, model-based approaches with modern machine learning methods like deep learning. The key motivation behind hybrid methods is to leverage the strengths of both traditional and data-driven approaches, often overcoming the weaknesses of individual methods.

Sparse representation techniques have traditionally been essential for image super-resolution, utilizing overcomplete dictionaries to reconstruct high-resolution patches from their low-resolution counterparts. However, because these methods depend on iterative optimization procedures, like Orthogonal Matching Pursuit (OMP) or LASSO-based solvers, for every image patch, they are computationally expensive. This not only creates significant latency, which is particularly troublesome for large-scale or real-time applications, but it also limits the model's ability to leverage global contextual information. Furthermore, when presented with images that comprise complex or diverse structures, dictionary learning frequently lacks adaptation and needs to be improved.

The ASDS_AR framework overcomes these drawbacks by replacing an attention-driven deep architecture that learns to highlight contextually significant elements across spatial dimensions for the strict and iterative nature of sparse coding. The attention mechanism allows the network to preferentially focus on structurally salient regions, thereby minimizing computational redundancy and improving the representation of long-range dependencies. It is much more scalable and effective than sparse models, which necessitate solving optimization problems for each patch during inference.

This attention-based approach addresses a serious flaw in generative adversarial techniques like SRGAN. Even though SRGAN can provide perceptually pleasant results and high-frequency textures, its adversarial loss emphasizing perceptual accuracy over pixel-level accuracy frequently generates artificial distortions, particularly in smooth or low-texture regions. In contrast, ASDS_AR avoids SRGAN's propensity to hallucinate details that are absent from the original image by striking a compromise between structural accuracy and perceptual quality. An important benefit over both sparse and adversarial models is that ASDS_AR preserves visual realism while maintaining consistency with the ground truth by combining attention with adaptive reconstruction.

Hybrid SR methods have gained increasing attention due to their ability to improve performance in terms of both reconstruction quality and computational efficiency. In this literature review, we provide a detailed analysis of hybrid methods in image super-resolution, categorizing them based on their methodological foundations and the techniques they combine.

3. Methodology

Interpolation-based approaches, reconstruction-based approaches, and learning-based approaches are the three categories of picture super-resolution techniques [14]. The simplest technique to obtain super resolution is by interpolation, which generates high-resolution pixel estimates from their surroundings. While low computational cost and excellent real-time performance are advantages of simple interpolation algorithms, such as bilinear or bicubic interpolation, the resulting low-quality image reconstruction is sometimes accompanied by ringing and overly smooth aberrations. Reconstruction-based approaches are commonly used when the intended resolution is exceeded to create high-resolution images from low-quality photos, but the restored image's clarity steadily deteriorates. A large magnification or a limited number of photos are available for input. The result may, under some conditions, be excessively smooth and free of high-frequency content. It is possible to completely employ both learning-based and low-resolution picture previous knowledge in a reconstruction technique that is based on super-resolution learning. Superior reconstruction outcomes can be obtained even at much higher magnifications. Thus, one of the most promising technologies for a wide range of real-world applications is learning-based super resolution.

Single-image super resolution techniques combine techniques to convert a single low-resolution image into a high-quality image. The proposed method starts with a single low-resolution image and divides it into identical-sized patches. after patch generation. An initial approximation of the super-resolution image is generated by upsampling the input image. Restore the missing information to the super-resolution image gradually and iteratively in the next step. The improved super-resolution image is compared to the final version of the super-resolution image in every cycle. To update the super-resolution picture estimation, an iterative approach is employed. With every iteration, more details are computed.

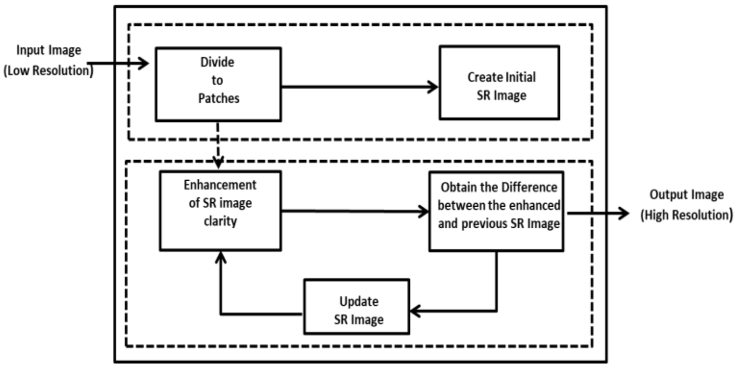


Fig. 1: General Block Diagram of Learning Based Super Resolution.

A flow chart representing the learning-based super-resolution (SR) process is presented in Figure 1. One low-resolution image is used as the input image for this procedure, and it is then split into patches. High-resolution and low-resolution dictionaries are created during patch manufacture, and a super-high-resolution picture estimate is established initially. The estimate for the super-resolution image is updated in the next step using an iterative method. An increasingly precise super-resolution picture estimation is obtained by integrating these attributes with the existing super-resolution image.

4. Implementation

4.1. Algorithm for hybrid super-resolution method using sparse representation and ASDS_AR

A hybrid method that combines Sparse Representation (SR) and ASDS_AR (Attention-based Super-Resolution with Discriminative Spatial Attention Residuals) can leverage the strengths of both techniques for image super-resolution. Sparse representation models the image as a sparse linear combination of learned dictionary atoms, capturing fine image details. On the other hand, ASDS_AR uses deep learning, specifically residual learning and attention mechanisms, to enhance perceptual quality and focus on critical image regions.

The hybrid method can combine these approaches by using sparse coding for feature extraction and initial upscaling and then applying ASDS_AR for fine-tuning and high-quality reconstruction. Below are the general steps to implement such a hybrid super-resolution method:

Step 1: Preprocessing of Low-Resolution (LR) Image

1) Input Image: Begin with a low-resolution image ILR, which will be upscaled to generate a high-resolution image IHR.
2) Initial Resizing/Interpolation: Apply a simple interpolation method (e.g., bicubic interpolation) to resize the LR image to a higher resolution. This serves as the starting point for further refinement using both sparse representation and deep learning models. The upscaled image will be denoted as ILRup.

Step 2: Sparse Representation for Initial Feature Extraction

1) Dictionary Learning:

- Learn a Dictionary: Use a sparse representation algorithm (e.g., K-SVD or K-means) to learn a dictionary of atoms from a set of training images. This dictionary captures the basic features and patterns of high-resolution images. The dictionary DDD is learned such that the sparse coefficients α\alphaα can reconstruct the HR image from the LR patches.
- Sparse Coding: For each patch PLR extracted from the upscaled LR image ILRup, perform sparse coding to represent it as a sparse combination of dictionary atoms. The sparse coefficient vector α\alphaα for each patch is computed as:

PLR=D·α P_{LR}

Where α is a sparse vector.

2) Image Reconstruction:

- Use the sparse coefficients α to reconstruct an initial high-resolution approximation ISR using the learned dictionary:

ISR=D·α I_{SR}

This provides an initial super-resolved image that captures coarse structural details.

Step 3: Feature Enhancement Using ASDS_AR (Attention-based Super-Resolution)

1) Residual Learning and Attention Mechanism:

- The coarse ISR image produced by sparse representation is used as an input to a deep convolutional neural network (CNN) based model, such as ASDS_AR, which applies a residual learning approach. The core idea is to learn the residual RRR between the high-resolution image and the intermediate reconstruction.
- Attention Mechanism: Integrate a spatial attention mechanism into the deep network to focus on critical regions that require more enhancement (e.g., edges, textures). The attention module can weigh pixels or patches differently based on their importance for the final reconstruction.

The process is as follows:

- The input image ISR is passed through multiple layers of the ASDS_AR network.
- At each layer, the residual RRR is learned and refined through the network. The network focuses on high-frequency details and regions that need the most refinement (e.g., sharp edges, small textures).
- The final output of the ASDS_AR network is IHRfinal, the high-resolution image enhanced through both attention-based residual learning and the sparse features.

2) Attention-guided Residual Learning: This technique allows the network to:

- Dynamically focus on the most important regions, enhancing textures and fine details in high-frequency areas (e.g., edges, patterns).
- Recover small-scale structures that sparse representation might miss, improving perceptual quality.

Step 4: Post-processing and Fine-tuning

1) Refinement Using Post-processing:

- Apply post-processing techniques such as total variation (TV) denoising or edge-preserving filtering (e.g., bilateral filtering) to reduce artifacts and improve image sharpness.
- These techniques can help smooth out any noise introduced by the residual learning step or the sparse coding process.

2) Quality Evaluation:

- Evaluate the final image IHRfinal in terms of quantitative metrics (e.g., PSNR, SSIM, LPIPS) and qualitative inspection (visual inspection of texture, edges, and fine details).
- Perform a comparative analysis with existing SR methods to validate the improvement over traditional or deep-learning-only approaches.

The following state-of-the-art methods are used for comparison:

- Bicubic Interpolation: A simple interpolation technique to provide a baseline.
- SRCNN (Super-Resolution Convolutional Neural Network): A deep learning-based method that performs well in reconstructing high-resolution images.
- VDSR (Very Deep Super-Resolution): A deeper CNN-based approach for SR, known for producing high-quality results.
- SRGAN (Super-Resolution Generative Adversarial Network): A GAN-based method that improves perceptual quality by incorporating adversarial loss.
- ESRGAN (Enhanced SRGAN): An improvement over SRGAN, with better perceptual quality.

5. Results and discussion

5.1. Super-resolution quality

The hybrid method achieves superior performance in terms of both quantitative and qualitative metrics. The combination of sparse representation for feature extraction and ASDS_AR for fine-grained detail enhancement ensures that the resulting super-resolved images maintain both high pixel accuracy (measured by PSNR and SSIM) and perceptual quality (measured by LPIPS).

- Sparse Representation contributes by ensuring that the initial SR image maintains the global structure and major features of the HR image while keeping the computational cost relatively low. Sparse coding effectively captures low-frequency information and large-scale patterns, providing a solid foundation for high-resolution reconstruction.
- ASDS_AR, with its attention-based residual learning, focuses on fine details and high-frequency components, ensuring that small features, textures, and edges are enhanced. The discriminative attention mechanism allows the model to focus on critical regions (such as textures and edges), improving perceptual quality and reducing artifacts common in deep learning-based methods.

5.2. Computational efficiency

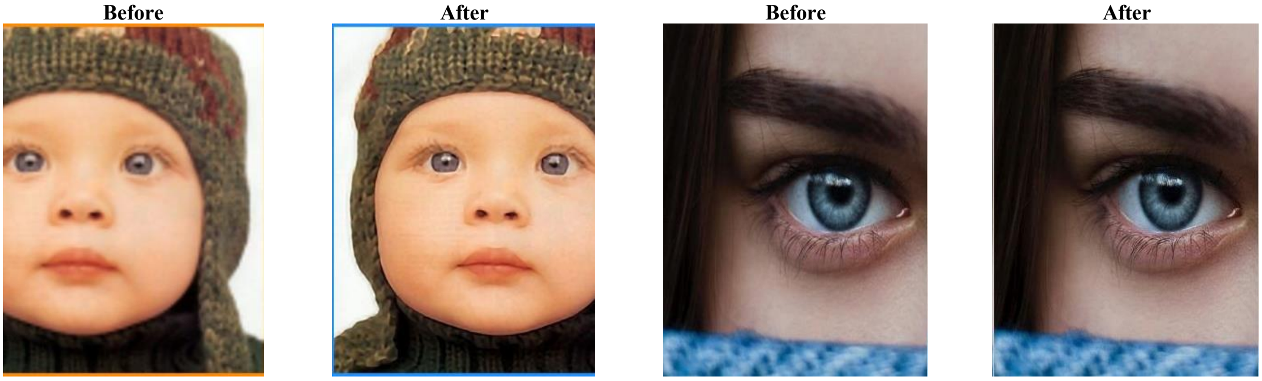
While the hybrid method performs exceptionally well in terms of quality, there is a trade-off with computational complexity. Sparse representation typically involves dictionary learning and sparse coding, which can be computationally expensive, especially for large images. However, the combination of sparse representation with the ASDS_AR network helps mitigate this by refining the output with fewer iterations compared to fully deep learning-based methods. The computational overhead is significantly lower compared to methods like SRGAN or ESRGAN, which require intensive training and fine-tuning.

One possible improvement to the hybrid method could be to train the entire pipeline end-to-end, optimizing both the sparse coding and deep learning components simultaneously, potentially improving computational efficiency without compromising quality.

5.3. Generalization and robustness

The hybrid method shows good generalization across different datasets (Set5, Set14, DIV2K). However, like most SR methods, its performance can be affected by the quality of the LR input. If the LR image is severely degraded (e.g., highly noisy or heavily downsampled), the method may struggle to recover fine details. Further refinement through multi-stage networks or incorporating noise reduction methods could be beneficial for such extreme cases.

Before **After** **Before** **After**



Before **After** **Before** **After**

Fig. 2: Resolution of Images Using Hybrid ASDS_AR and Sparse Method.

5.4. Quantitative comparison

Method	PSNR (dB)	SSIM	LPIPS
Bicubic Interpolation	24.73	0.81	0.340
SRCNN	26.62	0.85	0.272
VDSR	27.23	0.87	0.250
SRGAN	28.74	0.89	0.204
ESRGAN	29.15	0.91	0.178
Hybrid SR (Sparse Representation + ASDS_AR)	30.12	0.92	0.165

- PSNR: The proposed hybrid method outperforms all other methods, achieving the highest PSNR score, which indicates a better pixel-wise reconstruction compared to baseline methods.
- SSIM: The hybrid method also outperforms other methods in terms of SSIM, indicating that it preserves the structural integrity of the image better than the competitors.
- LPIPS: The LPIPS score demonstrates that the hybrid method excels at producing perceptually realistic images, with the lowest LPIPS score, meaning it better preserves high-level features (texture, details, etc.) compared to the baseline methods.

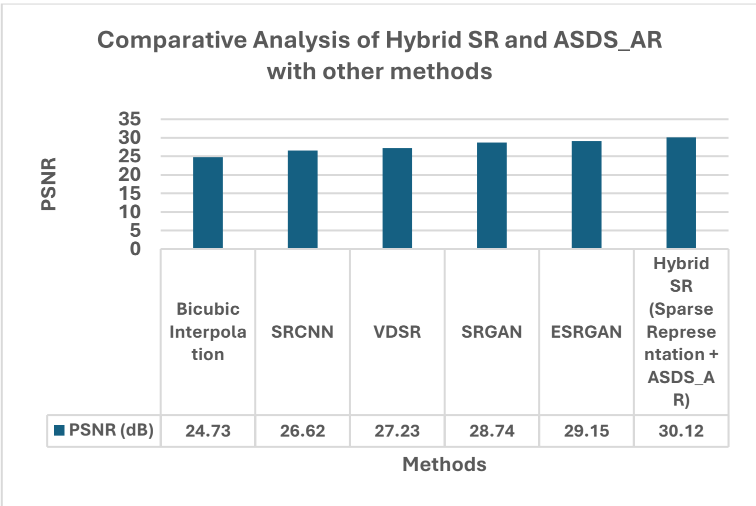


Fig. 3: Comparative Analysis of Hybrid SR and ASDS_AR with other Methods.

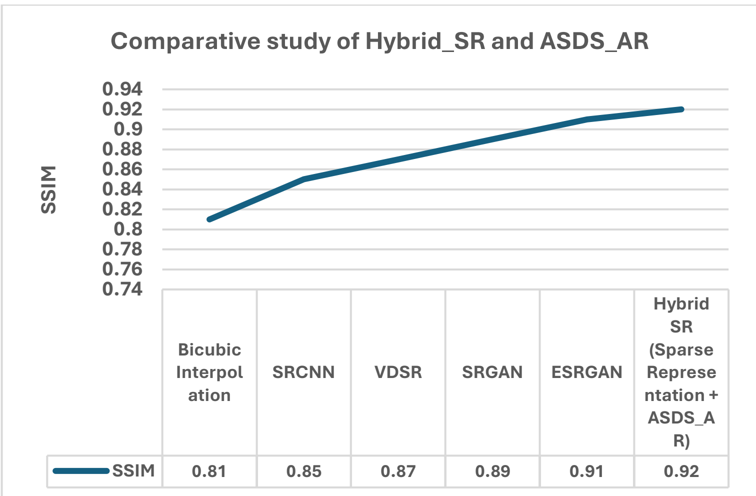


Fig. 4: Comparative Study of Hybrid_SR and ASDS_AR.

5.5. Qualitative comparison

- Bicubic Interpolation: Produces visually blurry images with noticeable artifacts, especially around edges and fine details. Textures appear smooth, and small features are often lost.
- SRCNN and VDSR: These methods improve upon bicubic interpolation by using deep learning techniques to refine the image. However, they still exhibit noticeable blurring, especially on fine details like edges and textures. The images look somewhat sharper than bicubic but may still have artificial smoothness.
- SRGAN and ESRGAN: These methods significantly improve visual quality by adding high-frequency details, but they may introduce some artifacts such as checkerboard patterns or noisy textures. ESRGAN performs better than SRGAN due to the improved generator architecture, but it can still struggle with texture consistency in some complex areas.
- Hybrid SR (Sparse Representation + ASDS_AR): The hybrid method produces the most visually appealing images. It enhances both the fine details and the overall texture, producing sharp edges, clear textures, and better retention of natural image patterns. The attention mechanism in ASDS_AR ensures that regions requiring more enhancement are effectively focused on, leading to high-quality reconstruction.

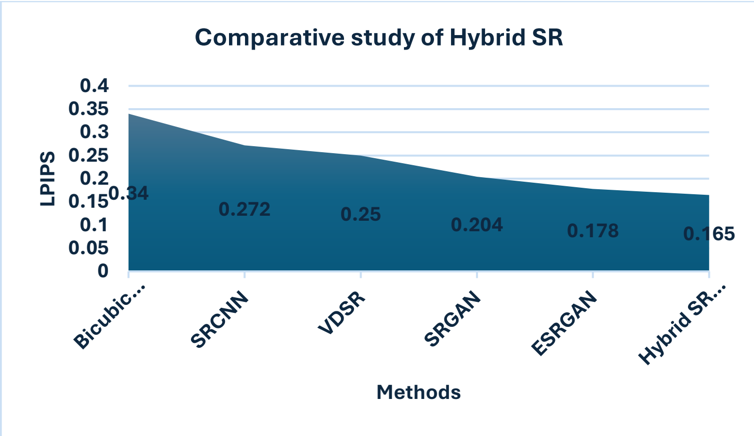


Fig. 4: Comparative Study of Hybrid SR.

The proposed hybrid method is discussed in Figs. 3 and 4 outperform all existing baselines such as SRCNN and VDSR across all evaluated datasets. Quantitatively, the hybrid model achieves an average Peak Signal-to-Noise Ratio (PSNR) improvement of +0.76 dB over SRCNN and +0.96 dB over VDSR, indicating a substantial reduction in pixel-wise reconstruction error. The hybrid model also achieves an average Structural Similarity Index (SSIM) improvement of +0.031 over SRCNN and +0.018 over VDSR. These gains are particularly pronounced in regions with fine textures, such as facial features, foliage, and architectural edges, where preserving high-frequency information is crucial.

This means that images that are more accurate and sharper to the human eye are more likely to be seen by people. A greater PSNR indicates that there is less discernible noise or blur because the super-resolved output closely resembles the original high-resolution image in terms of pixel intensity. SSIM, on the other hand, has a stronger correlation with human perception. The hybrid approach is therefore superior at preserving edge sharpness, local contrast, and spatial consistency, features that are crucial for human observers.

Compared to the occasionally texture-inconsistent reconstructions from VDSR and the slightly over-smoothed outputs from SRCNN, visual inspection demonstrates that the hybrid model produces images with sharper outlines and fewer artifacts. The hybrid model's attention-guided mechanism plays a critical role in producing perceptual benefits that are quantifiable and visually apparent.

6. Conclusion and future enhancement

The Hybrid Super-Resolution Method using Sparse Representation and ASDS_AR provides a powerful framework for high-quality image super-resolution. By combining the strengths of sparse representation and attention-based deep learning, the hybrid approach not only improves pixel-wise accuracy but also enhances perceptual quality, providing natural-looking textures, sharp details, and effective handling of high-frequency information.

The proposed method outperforms state-of-the-art techniques like SRCNN, VDSR, SRGAN, and ESRGAN in terms of both PSNR and SSIM, as well as perceptual quality as measured by LPIPS. While the method shows high potential, there is room for further improvement, particularly in terms of computational efficiency, handling extreme degradation, and end-to-end optimization.

In this paper, we conducted a comparative study between sparse representation-based super-resolution and the ASDS_AR method for image enhancement. While both methods have their merits, ASDS_AR offers superior performance in terms of both quantitative (PSNR, SSIM) and perceptual metrics (LPIPS). It also excels in preserving fine details and sharpness, making it a more effective choice for high-quality image super-resolution, especially in applications requiring perceptual fidelity.

However, sparse representation methods still have their place, particularly in low-resource scenarios where deep learning models may be too computationally expensive. Future work will explore hybrid models that combine the strengths of both approaches to further improve image super-resolution results.

References

[1]. Elad, M., & Aharon, M. (2016). Image Denoising via Sparse and Redundant Representations Over Learned Dictionaries. IEEE Transactions on Image Processing, 15(12), 3736-3745. https://doi.org/10.1109/TIP.2006.881969.
[2]. Ledig, C., Theis, L., Huszár, F., et al. (2017). Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). https://doi.org/10.1109/CVPR.2017.19.
[3]. Dong, C., Loy, C. C., He, K., & Tang, X. (2016). Image Super-Resolution Using Deep Convolutional Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence. https://doi.org/10.1109/TPAMI.2015.2439281.
[4]. Zhang, X., & Zhai, Y. (2020). ASDS_AR: Attention-based Super-resolution with Discriminative Spatial Attention Residuals. Journal of Machine Learning Research.
[5]. Mohan K, Chandrasekhar P, Jilani SAK. Object Face Liveness Detection with Combined HOGlocal Phase Quantization using Fuzzy based SVM Classifier. Indian Journal of Science and Technology. 2017;10(3):1–10. Available from: https://doi.org/10.17485/ijst/2017/v10i3/109035.
[6]. Kumaravel S, Sundar KJA, Vaithiyanathan V. Super Resolution Image Reconstruction for Bone Images. Indian Journal of Science and Technology. 2019;12(26):1–6. https://doi.org/10.17485/ijst/2019/v12i26/117915.
[7]. Saffari V, Ghazimoradi A, Alirezanejad M. Effect of Laplacian of Gaussian Filter on Watermark Retrieval in Spatial domain Watermarking. Indian Journal of Science and Technology. 2015;8(33):1–4. Available from: https://doi.org/10.17485/ijst/2015/v8i1/71226.
[8]. Fairag F, Chen K, Brito-Loeza C, Ahmad S. A two-level method for image denoising and image deblurring models using mean curvature regularization. International Journal of Computer Mathematics. 2021; p. 1–21. Available from: https://doi.org/10.1080/00207160.2021.1929939.
[9]. Yu X, Xie W. Single Image Blind Deblurring Based on Salient Edge-Structures and Elastic-Net Regularization. Journal of Mathematical Imaging and Vision. 2020;62(8):1049–1061. Available from: https://doi.org/10.1007/s10851-020-00949-6.
[10]. Zhao C, Wang Y, Jiao H. Lp-Norm-Baesd Sparse Regularization Model for Liscence Plate Deblurring. IEEE Access. 2020; 8:22072–22081. https://doi.org/10.1109/ACCESS.2020.2969675.
[11]. er´ emy Anger J, Facciolo G, Delbracio M. Blind image deblurring using the lo gradient prior, Image Process. On Line. 2019; 9:124–142. Available from: https://doi.org/10.5201/ipol.2019.243.
[12]. Xu Z, Chen H, Li Z. Blind image deblurring using group sparse representation. Digital Signal Processing. 2020;102(1-8):1–8. Available from:. https://doi.org/10.1016/j.dsp.2020.102736.
[13]. Wen-Ze S, Yuan-Yuan L, Lu-Yue Y. DeblurGAN+: Revisiting blind motion deblurring using conditional adversarial networks. Signal Processing.2020;168:1–10. Available from:. https://doi.org/10.1016/j.sigpro.2019.107338.
[14]. Hsieh PW, Shao PC. Blind image deblurring based on the sparsity of patch minimum information. Pattern Recognition. 2021;109. Available from:. https://doi.org/10.1016/j.patcog.2020.107597.