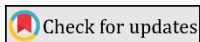


Modified very deep super resolution network - Gaussian mixture models based single image super-resolution: A comprehensive comparative analysis

Mehmet Akif Günen*¹ 

¹ Department of Geomatics Engineering, Faculty of Engineering and Natural Sciences, Gümüşhane University, 29100 Gümüşhane, Türkiye; akif@gumushane.edu.tr



Article History:

Received: 25 November 2025

Revised: 4 January 2026

Accepted: 28 January 2026

Published: 3 July 2026



Copyright: © 2026 by the author.

This article is an open access article distributed under terms and conditions of the Creative Commons Attribution (CC BY-SA) license. <https://creativecommons.org/licenses/by-sa/4.0/>

Abstract: Image Super-Resolution (SISR) is a vital, cost-effective solution for enhancing the spatial resolution of satellite imagery. However, a significant challenge in current SISR techniques is that while deep learning models excel at structural learning, they often struggle to accurately reconstruct intricate high-frequency details and preserve the statistical texture properties of high-resolution images. To address this gap, this paper proposes a novel hybrid method, termed mVDSR-GMM, which leverages the complementary strengths of Gaussian Mixture Models (GMM) and a modified Very Deep Super-Resolution (mVDSR) network. GMM is utilized to model the underlying statistical and probabilistic properties of the image for accurate high-frequency reconstruction while the mVDSR network enhanced with 25 convolutional layers and Leaky ReLU activations boosts the model's structural feature extraction and non-linear mapping capabilities. This hybrid approach technically integrates the mVDSR and GMM through a parallel merging of their estimated residuals within the luminance channel, thereby enhancing the spatial reconstruction of satellite imagery. The enhanced mVDSR network, with increased depth and Leaky ReLU activation functions, further boosts the model's learning capability and robustness. Through extensive evaluations using satellite imagery datasets and metrics such as PSNR, SSIM, BRISQUE, MSE, UIQI, SAM, and FD, mVDSR-GMM consistently outperformed existing techniques, demonstrating its effectiveness in improving spatial resolution. Experimental results indicate that mVDSR-GMM outperforms main competitors such as SwinIR and SinSR across all evaluation metrics, providing a robust reconstruction performance that is further validated by improved accuracy in the binary separation of urban structures (buildings and roads) from natural environments. The proposed mVDSR-GMM emerges as a promising tool for advancing satellite imagery resolution, with potential applications in remote sensing and geospatial analysis.

Keywords: Single image super-resolution; satellite imagery; Gaussian mixture models; very deep super-resolution network

Citation: Günen, M. A. (2026). Modified very deep super resolution network - Gaussian mixture models based single image super-resolution: A comprehensive comparative analysis. *Turk. J. Remote Sens.*, 8, 1829904. <https://doi.org/10.51489/tuzal.1829904>

1. Introduction

Satellite imagery assumes a critical role as a primary geoinformation source, facilitating various applications across Earth Science disciplines, such as environmental monitoring (Günen, 2022), crop identification (Dai et al., 2020), land use/cover mapping (Arabi Aliabad et al., 2023), disaster management (Atasever & Günen, 2021), decision support systems (Şan et al., 2021), and climate modeling (Leinonen et al., 2020). However, Earth observation satellites often struggle to capture both high spatial and spectral data due to inherent constraints in optic and sensor technology, as well as the high cost of such hardware capabilities. This fundamental trade-off significantly limits the potential of satellite imagery for applications requiring fine-grained detail. While multi-frame super-resolution (SR) offers a theoretical opportunity to amalgamate information from multiple low-resolution (LR) images, its real-world efficacy is constrained by recording precision, noise, and the impracticality of acquiring diverse multitemporal data for the same scene. Consequently, Single Image Super-Resolution (SISR) through software-based algorithmic approaches

emerges as a vital, cost-effective solution for eliminating these hardware limitations and enhancing the spatial resolution of satellite data.

The primary goal of image SR is to transform LR images characterized by blurriness and indistinct details into clear, high-resolution (HR) counterparts, thereby enhancing visual perception and overall quality. This task is mathematically challenging because multiple HR solutions can exist for a single LR input. In the scientific community, SR is frequently discussed using terms such as interpolation, image scaling, and upscaling. Beyond remote sensing (Wang et al., 2022a), the impact of SR extends to small object detection (Courtrai et al., 2020), face recognition (Farooq et al., 2021), medical diagnostics (Özyurt et al., 2020), astronomy (Dabbech et al., 2022), and forensics (Salguero-Cruz et al., 2022).

Methodologically, image SR techniques are categorized into interpolation-based, reconstruction-based, transformer-based, and learning-based methods (Lepcha et al., 2023). Interpolation methods are valued for their simplicity and computational efficiency (Singh & Singh, 2020), yet they often fail to capture intricate details. Reconstruction-based approaches preserve overall structure but struggle with noise and diverse content (Chen et al., 2021; He et al., 2023). Modern Transformer-based methods excel at capturing long-range dependencies but face challenges regarding high computing demands and memory storage (Lei et al., 2021; Cheng et al., 2023). Learning-based techniques, particularly those rooted in Deep Neural Networks (DNNs), have significantly enhanced SR performance by automatically learning complex textures and features from input data, though they require meticulous tuning to avoid overfitting (Jiang et al., 2021; Wang et al., 2022b).

Recent advancements have integrated residual learning into deep recursive neural networks to improve reconstruction quality. Various architectures have been proposed, including Convolutional Neural Networks (CNN) using ParametricReLU (Joshi et al., 2022), Generative Adversarial Networks (SRGAN) utilizing perceptual loss (Ledig et al., 2017), and end-to-end mappings like SRDCN, DSRDCN, and ESRDCN (Ran et al., 2020). Other notable methods include the Super Resolution Convolutional Neural Network (SRCNN) (Dong et al., 2014), the Residual Dense Network (RDN) (Zhang et al., 2018), and techniques utilizing patch self-similarity (MHSR) or sparse coding (SCSR) (Chen & Fowler, 2012; Glasner et al., 2009; Yang et al., 2010). Specifically, (Kim et al., 2016) introduced the Very Deep Super-Resolution (VDSR) network inspired by VGG-net, while (Sandeep & Jacob, 2016) proposed using Gaussian Mixture Models (GMM) trained on patch pairs to enable HR inference.

In this study, a novel hybrid structure is proposed. The method combines a modified VDSR model (mVDSR) with a GMM for SISR. The method uses residuals from both the GMM and mVDSR models to enhance the image SR. QuickBird-2, Ikonos-2, and GeoEye-1 test images have been used to evaluate performance. As can be seen from the literature review, a large number of image SR methods have been proposed. Among these methods, Bicubic, DSRDCN, ESRDCN, MHSR, SCSR, SRDCN, SwinIR, SinSR, and Real-ESRGAN methods have been preferred for comparison. Subsequently, qualitative and quantitative analyses have been conducted. Finally, the images produced by these methods were then used to distinguish buildings and roads from sea and green areas using the histogram curve-fitting threshold method. The main contributions of this work to the literature are as follows.

- 1) A novel hybrid model for SISR is proposed, combining an mVDSR architecture with a GMM that has not previously been investigated in the literature.
- 2) The mVDSR enhances the feature extraction capability of the original network by employing a total of 25 convolutional layers, integrated with 3 Leaky ReLU and 21 ReLU layers to improve non-linear mapping and learning efficiency.
- 3) The combination of residual images from mVDSR and GMM results in more accurate and detailed image reconstruction, outperforming previous approaches.
- 4) The suggested technique is carefully tested against many cutting-edge algorithms utilizing satellite images, with evaluations based on Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), Blind/Referenceless Image Spatial Quality Evaluator

(BRISQUE), Mean Squared Error (MSE), Universal Image Quality Index (UIQI), Spectral Angle Mapper (SAM), and Fidelity-Deformation (FD) metrics.

5) The approach has been shown to be effective in the separation of urban structures from natural features, demonstrating its usefulness in remote sensing and urban planning applications.

2. Materials and Methods

The foundation of the mVDSR component within our hybrid framework relies on two distinct datasets to ensure a broad learning base. The first is the DIV2K dataset (Agustsson & Timofte, 2017), widely recognized in the super-resolution community for its high-quality diverse images. The second dataset comprises aerial photographs supplied by the Turkish General Directorate of Mapping, which are not publicly available. These photographs furnish the model with high-altitude spatial features pertinent to remote sensing. To facilitate efficient training, 200 images were randomly selected from DIV2K and 200 images were cropped from the aerial dataset into patches of 512x512. During the mVDSR training phase, these were further processed into 64x64 patches with no overlap to optimize the network's learning of localized textures. To enable the hybrid framework to reconstruct HR details from low-quality inputs, it is necessary to simulate the degradation process that occurs during satellite image acquisition. This process involves blurring and subsampling the original HR images to generate their LR counterparts. The random blur kernels are 11×11 anisotropic Gaussians, with lengths (λ_1, λ_2) independently distributed from a uniform range of u (0.6, 5), and a rotation angle θ from u $[-\pi, \pi]$. To further enhance the realism, uniform multiplicative noise is introduced to the blur kernel before normalizing its sum to one. To be used in the methods within the scope of the application, 200 images randomly selected from DIV2K and 200 images cropped from aerial photographs with dimensions of 512 x 512 were used. Images representing the datasets are presented in Figure 1.

A critical aspect of this study's novelty is the evaluation of the hybrid framework exclusively on satellite imagery that was not encountered during the training process. Satellite images from GeoEye-1, QuickBird-2, and Ikonos-2 were intentionally excluded from training due to their high acquisition costs and limited volume. This study demonstrates the effectiveness of the proposed method by evaluating it exclusively on satellite images that were not used during training. Testing solely on satellite imagery with varying spatial resolution highlights the method's ability to generalize across datasets and its practical applicability in scenarios with limited or unavailable training data. This approach represents a key contribution, showcasing the method's robustness in processing unseen satellite data. Table 1 presents the spatial resolutions of the satellite testing images, including multi-spectral (MS) and panchromatic (PAN) resolutions. The images were captured from the region of Zonguldak, Türkiye.

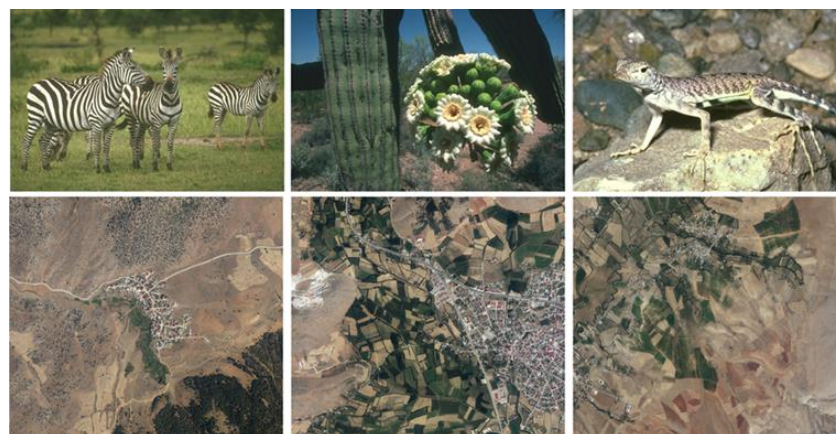


Figure 1. Overview of two datasets. The top row shows the DIV2K, the bottom row shows the aerial photo dataset

Table 1. Spatial resolution details of testing data

Product	Spatial Resolution (PAN/MS)
GeoEye-1	0.46/1.64 m
Ikonos-2	1.00/4.00 m
QuickBird-2	0.61/2.40 m

2.1. GMM-based image SR

GMM-based image SR uses GMM to enhance the resolution of images. GMMs are a statistical model that represents a mixture of multiple Gaussian distributions. In the context of image SR, GMMs can be employed to model the underlying statistical properties of HR images. GMMs offer a probabilistic representation, allowing a detailed modeling of pixel intensities and enhancing the overall representation of image characteristics. The GMM-based image SR method developed by (Ogawa et al., 2012a) and (Ogawa et al., 2012b) is used in this article. However, in this study, the luminance channel of the image was used. This approach was designed for GMM-based image conversion with the goal of transforming LR images into HR versions. The process involves creating a conversion function that establishes a relationship between the original image and its down-sampled counterpart. The Bicubic interpolation method was employed for both enlarging and reduction of images. The general workflow diagram of the proposed method is presented in Figure 2.

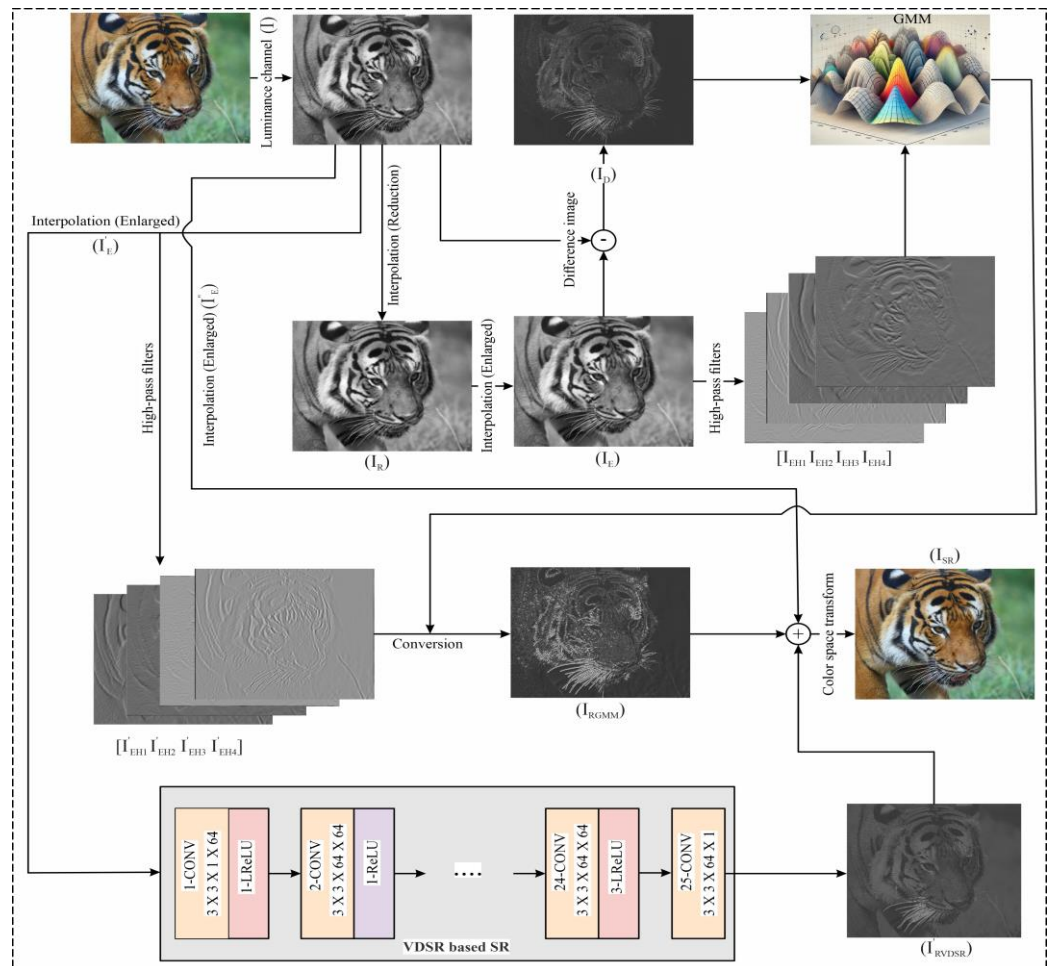


Figure 2. Flowchart of proposed method (I_D , I_{RGMM} , and I_{RVDSR} were automatically adjusted to increase visualization)

In GMM-based conversion, let the source and the target images shown as $x = [x_1, x_2, \dots, x_n]^T$ and $y = [y_1, y_2, \dots, y_n]^T$, respectively. GMM can be represented as

$P(z) = \sum_{m=1}^M \alpha_m N(z; \mu_m^z, \mu_m^z)$. Where joint distribution of vector is $z = [x^T, y^T]^T$, the weight of the m -th mixture is denoted as α_m , and M represents the total number of mixtures. Mean vector and variance-covariance matrix of m -th mixture trained by the expectation-maximization algorithm presented in Equation 1.

$$\Sigma_m^z = \begin{bmatrix} \sum_m^{xx} & \sum_m^{xy} \\ \sum_m^{yx} & \sum_m^{yy} \end{bmatrix}, \mu_m^z = \begin{bmatrix} \mu_m^x \\ \mu_m^y \end{bmatrix} \quad (1)$$

Subsequently, the transformation function from the source image to the target image is formulated based on the minimum Mean Square Error (MSE). Equation 2 is used to transform the source image into the target image.

$$y = \sum_{m=1}^M w_m(x) [\mu_m^y + \sum_m^{yx} (\sum_m^{xx})^{-1} (x - \mu_m^x)], w_m(x) = \frac{\alpha_m N(x; \mu_m^x, \sum_m^{xx})}{\sum_{j=1}^M \alpha_j N(x; \mu_j^x, \sum_j^{xx})} \quad (2)$$

Learning and estimation steps were followed to perform the GMM-based image SR process. During the learning phase, the conversion function is built from the input image and its self-reduction image using a GMM. In GMM training, the output data used as residual, that is, the difference image (I_D), is obtained from the difference between the is luminance channel of original input image (I) and the enlarged image (I_E) of the reduced image (I_R) from I . The input data for GMM is obtained by extracting high frequency components from IE (I_{EH1} , I_{EH2} , I_{EH3} , and I_{EH4}). These components are obtained with the by filtering the I_E using kernels presented in Equation 3.

$$\begin{aligned} H_1 &= [0 \ 0 \ 1 \ 0 \ 0 \ -1] \\ H_2 &= [0 \ 0 \ 1 \ 0 \ 0 \ -1]^T \\ H_3 &= [1/2 \ 0 \ 0 \ -1 \ 0 \ 0 \ 1/2] \\ H_4 &= [1/2 \ 0 \ 0 \ -1 \ 0 \ 0 \ 1/2]^T \end{aligned} \quad (3)$$

With the features obtained with Equation 3, I_D is decomposed into small image patches. Where k is the patch index, the k th image patch production is shown by Equation 4.

$$\begin{aligned} x_k &= [I_{EH1}^T \ I_{EH2}^T \ I_{EH3}^T \ I_{EH4}^T]^T \\ y_k &= [I_{D_k}]^T \end{aligned} \quad (4)$$

In the estimation phase, the conversion function is used to restore the residual image (I_{RGMM}). In this context, high-frequency components (I'_{EH1} , I'_{EH2} , I'_{EH3} , and I'_{EH4}) are obtained after enlarging (I_E) of I . The acquired images are broken down into patches. Applying the transformation function outlined in Equation 2 to these patches yields the I_{RGMM} from the trained GMM model.

2.2. m-VDSR

The VDSR method represents one of the image SR techniques published at CVPR in 2016 (Kim et al., 2016). This method, inspired by the VGG-net, aims to obtain HR images using DNN, specifically, multi-layered CNN. The network takes a LR image and transforms it into HR image. Throughout this process, the network systematically examines the image, progressively adding more details at each stage. The depth of the network (number of layers)

is a defining characteristic of VDSR. The original VDSR features a deep structure with 41 layers, incorporating 20 convolution layers, 19 rectified linear unit (ReLU) layers, input and regression output layers. Each of the convolution layers is equipped with 64 filters of dimensions 3×3. In the training process of m-VDSR, the objective is to estimate the data derived from the difference between I_{HR} and I'_E as a residual image, as described in Equation 5.

$$I_{RVDSR} = I_{HR} \ominus I'_E \quad (5)$$

In Equation 5, $\ominus$ represents Hadamard subtraction operator and IRVDSR refers residual image, I'_E represents the luminance channel of enlarged LR image obtained through Bicubic interpolation, while I_{HR} signify the luminance channel of corresponding input HR image. In training step, I_{RVDSR} at different scale levels (x2, x3, and x4) are utilized as outputs, while the corresponding I'_E at different scale levels are used as inputs. Thus, at the end of the training, when test or validation images are given to the trained model, the corresponding residual image, I'_{RVDSR} , is obtained for the respective LR image.

In this study, the image SR performance was enhanced through modifications made to the VDSR (m-VDSR) method. Unlike the original VDSR, the enhanced model consists of 25 convolution layers, 3 Leaky ReLU, 21 ReLU, and includes input and regression output layers. Increasing the number of layers enables the network to learn more parameters, recognize more complex patterns, enhance the tendency to learn more general and diverse dataset patterns, become more invariant to changes like translation, scale, and rotation, and better learn a hierarchical representation of features from input to output. Leaky ReLU is incorporated into the network after the 1st, 13th, and 24th convolution layers to assist the model in gaining better learning capabilities and attempting to overcome the issue of dead neurons. In some cases, Leaky ReLU proves to be more resilient to noise compared to ReLU. The general representation of the training process of mVDSR is presented in Figure 3.

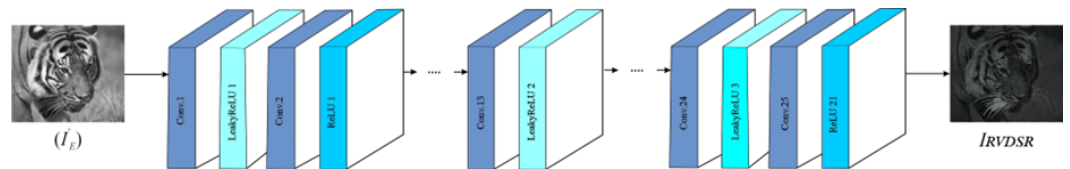


Figure 3. The general representation of the training of m-VDSR-based SISR (IRVDSR has been automatically adjusted to increase visualization)

To achieve SR image (I_{SR}) using the residuals derived from GMM-based image SR and m-VDSR, Equation 6 is used to obtain the new luminance channel (I_{NL}).

$$I_{NL} = I'_E \oplus (I_{RGMM} \oplus I'_{RVDSR}) \% 2 \quad (6)$$

In Equation 6, $\oplus$, $\%$ denotes Hadamard summation and division operators, respectively. In the YCbCr color space, Y, Cb, Cr represent luminance, blue, and red components of I'_E , respectively. I_{SR} is obtained using Equation 7.

$$I_{SR} = \text{cat}(I_{NL}, \text{Cb}, \text{Cr}) \quad (7)$$

2.3. Proposed method

The pixels located along the edges of the image have fewer neighboring pixels compared to those in the interior, presenting a challenge in accessing accurate information during the augmentation process. To ensure a fair comparison, pixels near the image boundary were cropped with three pixel before evaluation, following a similar approach as outlined in (Kim et al., 2016; Tai et al., 2017). The residual images were cropped before they were summation. The employed methodologies involve various hyperparameters, and their proper

configuration is crucial. In preparation for this study, different configurations, specifically 3, 6, 9, and 12 mixtures, were examined for GMM-based image SR for randomly selected 20 images from two dataset. It was established that utilizing 6 mixtures yielded optimal results. Additionally, a patch size of 3×3 was chosen for further implementation. In mVDSR training step, a patch size of 64×64 with no overlap was employed. The data augmentation process included random rotations in the [0, 90] range and reflection in the left-right direction. A learning rate of 0.1 was utilized with the stochastic gradient descent optimizer, incorporating a momentum of 0.85. Training was conducted with a batch size of 64, and the network trained 120 epochs. The implementation of a total of 850.5k learnable parameters took place on the MSI Delta 15 laptop, featuring an AMD Ryzen 9 9500 HX processor, RX6700 graphics card, and 32 GB RAM.

2.4. Performance assessment

As image SR techniques have become widely used, the evaluation of reconstructed images has become increasingly important. Image quality can be assessed subjectively and objectively. Subjective evaluation uses biased analysis to judge the naturalness and realism of an image based on human observers. This approach faithfully captures human perception. Conversely, objective evaluation uses mathematical algorithms and numerical computations to automatically assess the quality of an image (Günen, 2021; Civicioglu & Besdok, 2022; Pérez-Delgado & Günen, 2023; Civicioglu & Besdok, 2024). Objective quality assessment in this study involved the utilization of a comprehensive set of metrics, including PSNR (Chen et al., 2011), SSIM (Wang et al., 2004), BRISQUE (Mittal et al., 2011), MSE, UIQI, SAM, and the newly introduced FD metric.

PSNR is a popular metric for evaluating SR image quality by quantifying the relationship between signal's maximum potential energy and noise. SSIM, another key evaluation metric, measures both distortion and similarity between images using a perceptual model. Brisque determines image quality without reference images by modeling features using Gaussian distribution and analyzing contrast, brightness, and edge sharpness (Günen, 2021; Wang et al., 2022a; Lepcha et al., 2023; Al-Mekhlafi & Liu, 2024). The Fidelity-Deformation (FD) metric, introduced by (Günen et al., 2024), is a novel quality assessment approach for pan-sharpening methods that combines chromatic fidelity and spatial detail preservation. FD balances L0-Norm for chromatic fidelity with MSE for spatial accuracy, demonstrating the efficacy of L0-Norm pan-sharpening and establishing itself as a tool for analyzing SISR techniques. FD optimizes the quantity of pixels that maintain chromatic data for a given threshold value, T, within I_{SR} . The notation and formulation of the relevant metrics are presented in Table 2.

Table 2. Image quality assessment metrics

Metrics	Notation
MSE	$\frac{1}{m \times n} \sum_{i=1}^N (X_i - I_{SR_i})^2$
SSIM	$\left(\frac{2\mu_X \mu_{I_{SR}} + (0.01 \cdot \gamma)^2}{2\mu_X^2 \mu_{I_{SR}}^2 + (0.01 \cdot \gamma)^2} \right) \left(\frac{2\sigma_X \sigma_{I_{SR}} + (0.03 \cdot \gamma)^2}{\sigma_X^2 + \sigma_{I_{SR}}^2 + (0.03 \cdot \gamma)^2} \right) \left(\frac{\sigma_{X I_{SR}} + (0.03 \cdot \gamma)^2 / 2}{\sigma_X \sigma_{I_{SR}} + (0.03 \cdot \gamma)^2 / 2} \right)$
PSNR	$20 \log_{10} \left(\frac{p_{\max}^2}{\sqrt{MSE}} \right)$
FD	$FD = - \left(\frac{1}{K} \sum (I_{SR} - MSI \leq T) + \frac{1}{K} \sum (I_{SR} - PAN \leq T) \right) + MSE(I_{SR}, MSI) + MSE(I_{SR}, PAN), \quad T = 10$

*X: ground truth image, m, n: size of images, σ : variance, μ : mean, γ : dynamic range of the images, $p_{\max}$: maximum value of dynamic range, N: total number of bands

3. Results and Discussions

Assessing the impact of proposed modifications within an mVDSR is essential to comprehend their effectiveness. Figure 4 shows a comparison of the training loss curves between mVDSR and VDSR. This highlights the impact of modifications in mVDSR over baseline VDSR, illustrating the benefits of these changes. The importance of Figure 4 is underscored by its ability to quantitatively convey how the architectural enhancements introduced in mVDSR affect the training dynamics. This figure captures the reduction in training loss across 120 epochs, offering empirical validation of the proposed model's superiority regarding learning efficiency, stability, and optimization capabilities. In observing the training loss trajectories, it becomes evident that both mVDSR and VDSR manifest a rapid decrease in loss during the initial epochs, which is indicative of efficient feature learning processes. In particular, mVDSR persistently achieves lower loss metrics and attains stability faster than VDSR. The training durations for both the mVDSR and VDSR models, when applied to the aerial photograph dataset, were recorded at 20,487 and 19,450 seconds, respectively. This observation indicates that the mVDSR model required an additional 1,037 seconds, which is roughly 17 minutes, for its training process. mVDSR showed notable enhancements in reconstruction quality, with higher PSNR and SSIM metrics seen in Figure 5, despite a slight increase in training time. The enhanced performance significantly surpasses the modest additional computational cost. Figure 5 presents a visual and quantitative comparison of the SISR results obtained for two images, namely butterfly and panda, sourced from the DIV2K dataset, using mVDSR and VDSR. The findings clearly indicate that mVDSR surpasses VDSR in reconstruction quality, as evidenced by its higher PSNR and SSIM scores.

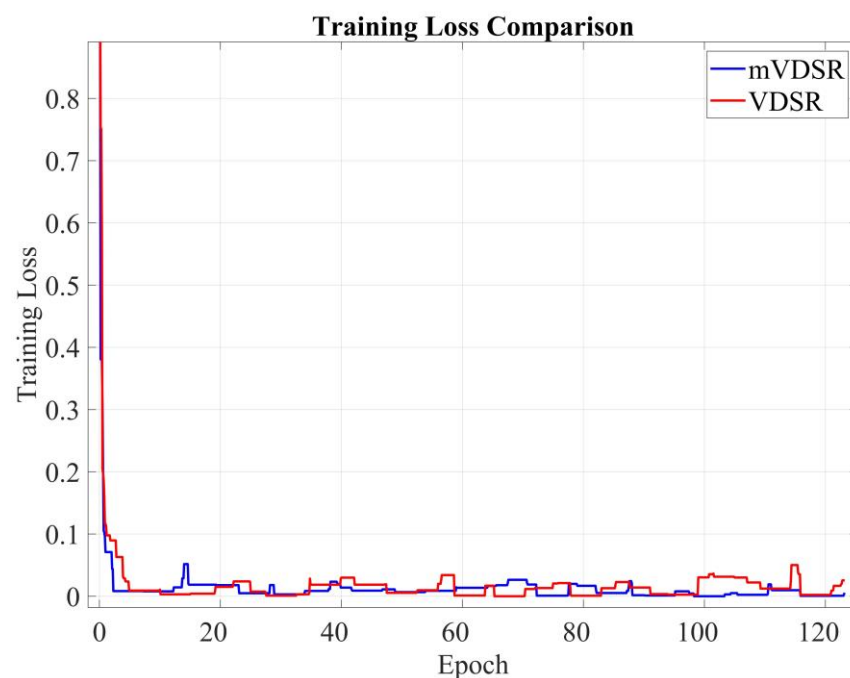


Figure 4. Training loss comparison of the proposed mVDSR and VDSR methods over 120 epochs. The blue curve represents mVDSR, and the red curve represents VDSR. The comparison highlights the performance improvement in terms of loss reduction during training

Images captured through geospatial imaging satellites were used for the performance analysis of the proposed method. Table 3-5 present a comparative analysis of performance metrics for different SR techniques applied to the GeoEye-1, QuickBird-2, and Ikonos-2 images, utilizing scaling factors of $\times 2$, $\times 3$, and $\times 4$. Table 3-5 show consistent patterns and nuanced variations in the efficacy of the method, demonstrating how different algorithms balance image clarity, and structural retention under varying conditions. High values of

PSNR, SSIM, and UIQI, along with low values of Brisque, MSE, FD, and SAM, collectively suggest a more favorable outcome in terms of success. For the $\times 2$ scaling factor, all approaches work well, but several stand out.

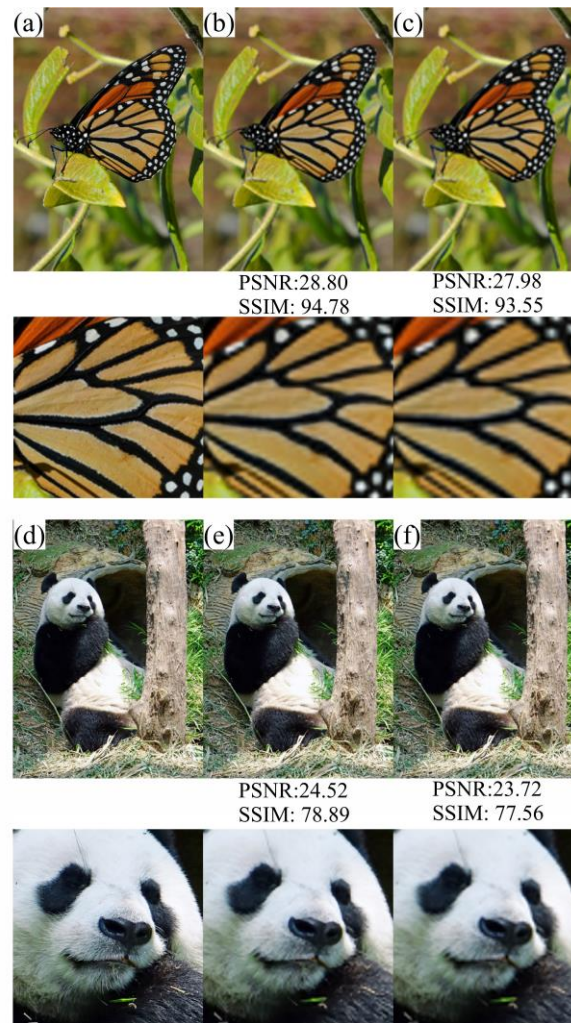


Figure 5. Comparison of HR and SISR results generated by different methods from DIV2K dataset, (a, d) HR reference images, (b, e) results obtained using the proposed mVDSR. (c, f), VDSR

The mVDSR-GMM method exhibits consistent superior performance across all metrics evaluated, with particular distinction in PSNR, SSIM, FD, and Brisque. Bicubic interpolation consistently demonstrates poorer performance across all evaluated metrics compared to deep learning-based methodologies, a result that aligns with expectations given its relatively simplistic nature. mVDSR-GMM shows notably higher PSNR, SSIM, and UIQI values, particularly at larger scales ($\times 3$ and $\times 4$), demonstrating enhanced capabilities in image reconstruction. SSIM and FD values are elevated in methods such as SinSR, SwinIR, and mVDSR-GMM, reflecting superior structural fidelity. Brisque scores are minimal for the likes of SinSR, SwinIR, and mVDSR-GMM, indicating heightened perceptual quality. With regard to MSE and SAM, methods such as mVDSR-GMM and SwinIR register the lowest values, indicating a more accurate preservation of image insights. As resolution increases (from a factor of $\times 2$ to $\times 4$), mVDSR-GMM consistently maintains outstanding performance levels, in contrast to Bicubic, which experiences a significant decline in quality. Notably, mVDSR-GMM surpasses alternatives such as SwinIR and SinSR in the quality metric SAM, affirming its superiority in both perceptual and spectral fidelity. SinSR significantly improves the perceptual quality of Ikonos-2 images, as evidenced by its lower BRISQUE scores compared to other methods, which highlights its dataset-specific efficacy. However, it still falls short in PSNR and SSIM, implying a trade-off between pixel accuracy and perceived visual quality. In

contrast, mVDSR-GMM exhibited lower performance with UIQI at x2. The DSRDCN approach performs modestly, particularly on QuickBird-2 images, but does not outperform mVDSR-GMM or SwinIR in overall metrics, showing that it is less capable of handling structural complexity and maintaining sharpness at this scaling factor.

Table 3. Statistical assessment of related methods for GeoEye-1 test images (×2, x3, and x4)

Image	x2												x3												x4											
	GeoEye-1																																			
	Method	PSNR	SSIM	Brisque	MSE	UIQI	SAM	FD	PSNR	SSIM	Brisque	MSE	UIQI	SAM	FD	PSNR	SSIM	Brisque	MSE	UIQI	SAM	FD	PSNR	SSIM	Brisque	MSE	UIQI	SAM	FD	PSNR	SSIM	Brisque	MSE	UIQI	SAM	FD
mVDSR-GMM	Real-ESRGAN	29.43	93.9	27.53	118.77	0.83	2.65	33.05	24.71	80.48	46.16	391.24	0.78	3.68	70.28	23.9	74.26	48.21	323.75	0.81	5.17	79.25														
	mVDSR	27.46	91.31	30.34	185.51	0.8	2.75	33.61	21.77	79.23	45.21	477.31	0.77	4.38	85.1	23.31	73.6	45.83	353.2	0.8	4.04	80.31														
	GMM	27.21	92.2	29.92	183.45	0.85	2.9	42.18	21.9	80.08	53.88	420.19	0.78	4.21	103.72	22.81	72.19	43.18	384.63	0.8	4.64	77.12														
	Real-ESRGAN	28.51	90.66	34.72	141.87	0.79	2.98	39.02	22.56	80.35	56.78	422.4	0.72	4.58	75.02	23.31	71.36	48.07	487.1	0.74	5	77.52														
	SinSR	29.37	92.36	27.25	116.92	0.88	2.95	31.05	24.2	75.45	50.38	399.24	0.76	3.99	83.69	23.34	70.5	48.6	485.28	0.78	4.54	82.8														
	SwinIR	29.35	91.17	30.13	134.5	0.86	2.91	59.09	23.86	76.21	46.38	388.84	0.74	4.2	89.12	23.37	69.09	52.94	506.87	0.79	4.42	79.66														
	SRDCN	29.28	91.82	28.86	124.96	0.81	3.22	86.42	20.59	65.13	42.12	834.75	0.76	3.7	83.46	22.61	72.95	44.76	373.43	0.77	5.5	83.32														
	SCSR	27.33	92.17	32.07	150.37	0.82	3.25	41.24	21.5	65.1	42.28	828.97	0.73	4.67	202.12	22.78	70.33	53.19	587.41	0.71	4.73	93.97														
	MHSR	27.54	92.93	25.32	141.19	0.8	3.11	48.79	21.61	65.48	48.32	806.96	0.71	3.97	182.15	22.75	69.44	49.56	512.47	0.74	5.78	82.7														
	ESRDCN	28.33	91.77	39.19	119.19	0.8	3.22	53.17	22.19	79.75	56.44	424.34	0.73	3.7	95.76	23.39	70.52	51.07	489.21	0.73	5.22	84.18														
	DSRDCN	26.88	88.17	33.05	157.25	0.79	3.12	56.02	21.2	65.64	45.77	805.2	0.77	3.67	339.17	22.25	65.67	53.07	584.85	0.72	4.95	89.88														
	BiCubic	26.87	87.6	40.6	193.75	0.78	3.33	80.01	20.59	64.68	42.13	818.18	0.77	4.87	182.34	22.01	64.2	52.41	595.13	0.73	6.89	92.13														

Table 4. Statistical assessment of related methods for QuickBird-2 test images (x2, x3, and x4)

Image	QuickBird-2												x2												x3												x4											
	mVDSR-GMM	mVDSR	GMM	Real-ESRGAN	SinSR	SwinIR	SRDCN	SCSR	MHSR	ESRDCN	DSRDCN	BiCubic	Method	PSNR	SSIM	Brisque	MSE	UIQI	SAM	FD	PSNR	SSIM	Brisque	MSE	UIQI	SAM	FD	PSNR	SSIM	Brisque	MSE	UIQI	SAM	FD														
	31.42	30.36	29.52	29.86	31.1	30.63	29.85	28.19	28.4	29.91	28.29	27.9																																				
	94.5	94.26	91.19	91.39	94.59	94.62	92.23	91.63	90.23	93.22	90.74	90.95																																				
	21.15	22.39	30.99	23.74	21.85	21.05	33.49	34.61	34.25	23.67	34.76	35.09																																				
	88.04	102.75	99.75	109.01	89.63	94.74	146.76	144.7	160.53	108.8	147.97	152.54																																				
	0.86	0.87	0.85	0.83	0.86	0.87	0.81	0.82	0.82	0.81	0.81	0.79																																				
	2.48	2.59	2.98	3.38	3.53	2.51	3.61	3.67	3.25	2.65	3.09	3.56																																				
	60.25	55.15	67.92	79.23	63.22	62.66	67.47	65.2	67.72	72.01	59.9	95.15																																				
	26.15	26.13	25.87	26.35	26.54	26.24	26.06	25.45	25.59	25.59	25.09	24.59																																				
	89.28	89.03	92.52	92.39	85.26	92.24	85.34	81.65	84.63	76.81	89.57	80.93																																				
	36.99	38.2	43.65	43.3	44.2	44.02	40.23	43.61	44.08	43.63	39.36	43.09																																				
	238.62	240.15	179.04	178.01	258.64	177.89	245.34	313.08	266.7	314.39	291.26	326.34																																				
	0.77	0.75	0.74	0.71	0.74	0.77	0.71	0.72	0.7	0.76	0.71	0.69																																				
	2.97	3.52	3.74	6.08	5.77	3.44	5.51	4.25	6.01	4.53	4.52	6.68																																				
	52.65	54.63	64.9	72.46	54.95	60.55	63.21	68.19	65.13	80.5	98.74	104.08																																				
	24.31	23.75	23.58	24.64	24.73	23.78	23.75	23.31	23.75	23.94	23.71	22.76																																				
	77.9	76.02	84.87	85.86	76.79	83.94	82.88	73.36	76.02	77.25	68.14	71.73																																				
	43.78	46.79	46.15	52.96	46.94	43.41	42.5	47.55	47.09	44.73	47.1	48.66																																				
	403.06	419.49	309.06	279.77	403.34	331.39	344.8	477.63	427.23	401.32	480.23	498.91																																				
	0.78	0.77	0.77	0.78	0.77	0.78	0.75	0.72	0.74	0.73	0.71	0.7																																				
	4.03	4.45	4.66	4.3	4.15	4.08	4.82	5.73	4.62	5.93	5.78	7.82																																				
	80.02	80.26	81.69	81.39	80.95	86	80.6	92.18	99.04	95.17	94.83	100.58																																				

Table 5. Statistical assessment of related methods for Ikonos-2 test images (×2, ×3, and ×4)

Image	Ikonos-2																x2	x3	x4																																																																																																																																																																																																																																																					
	Method	PSNR	SSIM	Brisque	MSE	UIQI	SAM	FD	PSNR	SSIM	Brisque	MSE	UIQI	SAM	FD	PSNR				SSIM	Brisque	MSE	UIQI	SAM	FD																																																																																																																																																																																																																																															
Ikonos-2	mVDSR-GMM	BiCubic	DSRDCN	ESRDCN	MHSR	SCSR	SRDCN	SwinIR	SinSR	Real-ESRGAN	GMM	mVDSR	28.89	27.04	28.28	27.25	27.94	27.93	27.63	26.7	26.08	27.72	26.39	25.63	93.48	88.58	96.03	91.55	95.86	91.67	92.21	91.87	94.24	91.69	94.95	87.91	19.72	32.62	29.92	20.31	29.94	20.4	20.23	32.06	24.58	20.56	32.26	30.74	161.74	230.96	122.35	189.73	120.37	163.92	172.13	233.86	183.48	183.67	196.39	258.95	0.84	0.81	0.82	0.88	0.82	0.83	0.81	0.81	0.8	0.8	0.81	0.79	2.98	2.99	2.74	3.01	2.97	2.95	3.49	3.2	3.3	3.03	3.46	3.78	44.78	48.14	40.5	57.85	40.12	45.53	53.98	72.17	54.23	54.42	57.43	75.85	24.15	23.7	23.91	24.28	24.6	24.04	23.78	23.13	23.13	24.36	23.32	22.64	78.58	78.88	82.62	81.18	81.63	77.84	80.65	77.5	77.31	81.86	83.71	76.6	35.99	41.09	40.82	41.44	40.88	32.34	41.16	43.3	31.38	41.45	44.26	44.5	303.44	284.59	395.86	422.08	404.83	317.02	426.19	493.99	322.48	396.92	495.72	514.75	0.79	0.78	0.78	0.74	0.79	0.76	0.76	0.76	0.74	0.74	0.73	0.72	4.94	4.6	4.56	4.87	4.39	4.8	4.89	6.04	5.01	5.91	5.8	6.34	92.12	90.36	105.75	101.25	103.63	95.25	110.63	150.96	100.25	145.45	157.2	200.91	22.54	21.52	21.91	22.52	22.43	22.71	21.85	21.91	21.81	21.66	21.74	20.98	72.93	77.78	78.53	71.16	79.16	72.02	68.36	67.38	75.05	77.85	67.77	66.76	45.06	42.2	43.61	52.32	49.26	47.89	51.34	59.52	51.33	42.12	51.72	54.34	637.7	509.86	466.07	656.98	432.69	642.03	735.77	760.41	738.59	494.17	738.12	759.95	0.75	0.75	0.73	0.7	0.74	0.75	0.74	0.73	0.73	0.74	0.74	0.73	6.68	4.89	4.93	5.32	4.95	4.27	4.68	5.29	6.02	6.09	5.22	6.68	131.28	132.36	129.9	135.25	128.42	150.98	195.36	198.2	200.12	150.25	205.15	215.54

As the scaling factor increases to ×3, differences between approaches become more noticeable. Overall performance drops, but mVDSR-GMM maintains strong PSNR, FD, and SSIM scores, particularly for GeoEye-1 and QuickBird-2 images. This resilience shows its ability to maintain quality reconstructions even as the task becomes harder. SwinIR performs competitively in Table 3-5, notably for QuickBird-2 images, where it achieves an MSE value superior to mVDSR-GMM. This suggests that while SwinIR may not equal mVDSR-GMM's structure retention (SSIM), it provides lower MSE, implying better balance of pixel-level precision and generalization. This indicates SwinIR could be viable for some image sets or when optimizing for pixel-wise error reduction. At this scaling factor, DSRDCN outperforms others in SSIM on Ikonos-2 images, demonstrating its strength in preserving structural elements at significant resolution reductions, attributed to its network architecture promoting hierarchical feature extraction. Despite lower PSNR, this SSIM performance makes it competitive for applications requiring structural precision.

At a highest scaling factor of $\times 4$, the performance disparity between approaches increases, as super-resolution reveals each algorithm's limitations. mVDSR-GMM outperforms competitors in PSNR, UIQI, and SSIM for GeoEye-1 images, showing reliability in high-demand settings. Even at this extreme scaling, it maintains maximum SSIM and *Brisque*, critical for high-resolution outputs from downscaled inputs. SwinIR and DSRDCN exhibit trade-offs at this level. SwinIR, despite lower SSIM scores than mVDSR-GMM, maintains low MSE and competitive *Brisque* values, indicating it produces pleasant results with smaller pixel errors. This balance could be useful where visual quality exceeds structural faithfulness. SinSR achieved a significant SSIM score of $\times 4$ for Ikonos-2 images, demonstrating that deep feature extraction can maintain geometry and structural information under challenging circumstances. This aligns with its prior performance, indicating promise for preserving structural elements in high-noise, high-scale environments. mVDSR-GMM consistently outperforms in PSNR and SSIM across all scaling factors and datasets, making it most reliable for pixel-level fidelity and structural preservation. Its performance on GeoEye-1 and QuickBird-2 images suggests its architecture suits satellite imaging requiring geometric and radiometric precision. SwinIR excels at $\times 3$ and $\times 4$, achieving competitive MSE and *Brisque* scores. This versatility suits applications where perceived quality exceeds pixel-level precision, like remote sensing visualization. The Bicubic interpolation approach performs consistently across the satellite images but produces the worst results.

It is very important in qualitative comparison as well as statistical comparison. In this regard, the scaled results $\times 2$, $\times 3$, and $\times 4$ for GeoEye-1 are presented in Figure 6, Figure 7, and Figure 8, respectively. The scaled results of $\times 2$, $\times 3$, and $\times 4$ for QuickBird-2 are presented in Figure 9, Figure 10, and Figure 11, respectively. The scaled results $\times 2$, $\times 3$, and $\times 4$ for Ikonos-2 are presented in Figure 12, Figure 13, and Figure 14, respectively. For a more precise evaluation of the performance disparities, visual results are best examined at high magnification where minor variations in edge sharpness and texture details are more noticeable. Visual interpretation provides a critical complement to numerical metrics by enabling the identification of complex patterns, fine textures, and structural features that standard algorithm may not fully represent. It has the potential to excel in understanding patterns, colors, and structures not captured by numerical metrics, which could enable a more comprehensive, human-centered image evaluation. As expected, the quality of the reconstructed images decreases as the scale factor increases. This is because image SR becomes more challenging with larger-scale gaps. In the qualitative comparison of image SR methods, it becomes easier for the human eye to distinguish different details as the scale size increases. There is still a noticeable absence of spatial and spectral details compared to the HR image. The numerical results are further supported by a visual comparison of the SR methods across various scaling factors. Sharper details and fewer artifacts were consistently produced in the images produced by the mVDSR-GMM method. Compared to alternative methods, mVDSR-GMM was able to preserve the structural details of the GeoEye-1 $\times 2$ image (Figure 6). Similar patterns were seen for the mVDSR-GMM-produced images that closely resembled the HR reference for the QuickBird-2 $\times 2$ image (Figure 9) and the Ikonos-2 $\times 2$ image (Figure 12). The visual quality of mVDSR-GMM remained consistent even when the scaling factor was increased to $\times 3$ and $\times 4$. For example, mVDSR-GMM preserved clarity and decreased blurring compared to other methods in the GeoEye-1 $\times 3$ image (Figure 7) and QuickBird-2 $\times 3$ image (Figure 10). Although all methods showed some degradation in image quality, mVDSR-GMM still produced superior visual results with well-preserved edges and textures when scaling up to $\times 4$ (Figure 8, Figure 11, and Figure 14). While GAN-based models achieve high perceptual quality, they often generate hallucinated high-frequency details that appear realistic but lack factual basis in ground truth. In remote sensing, these artifacts can lead to misinterpretation of small-scale features like building boundaries or land cover types. mVDSR-GMM prioritizes spectral and spatial fidelity over visual sharpness.

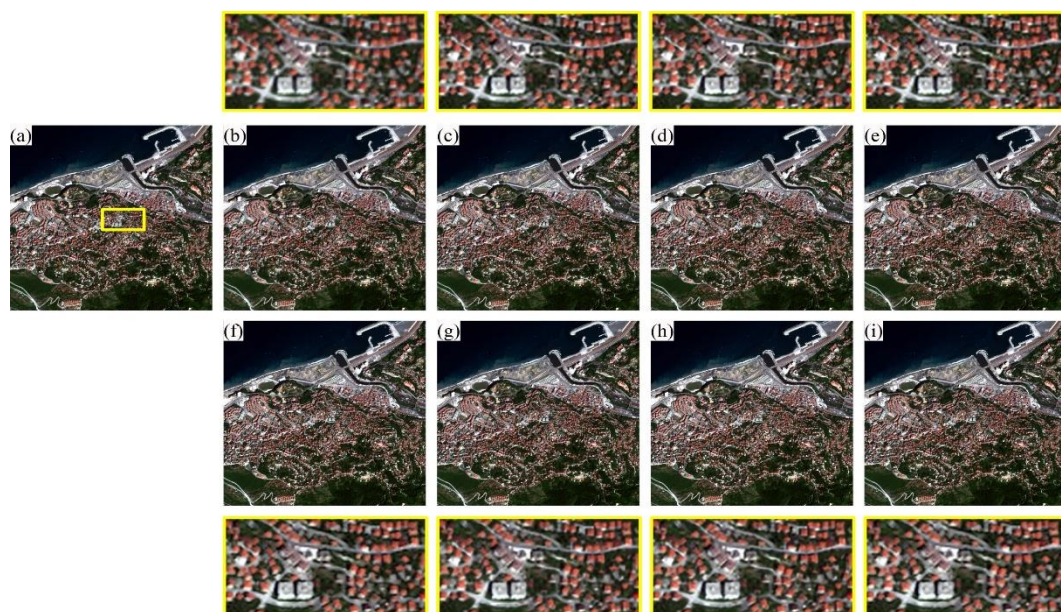


Figure 6. GeoEye-1 x2 scaled results (a) HR image, (b) Bicubic, (c) SwinIR, (d) DSRDCN, (e) SinSR, (f) Real-ESRGAN, (g) GMM, (h) mVDSR, and (i) mVDSR-GMM

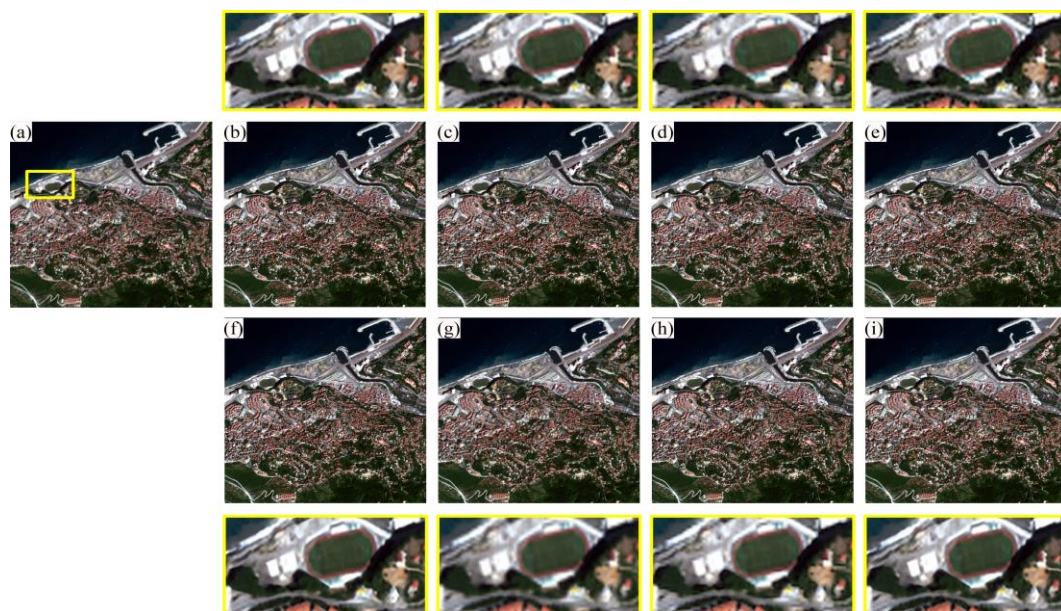


Figure 7. GeoEye-1 x3 scaled results (a) HR image, (b) Bicubic, (c) SwinIR, (d) DSRDCN, (e) SinSR, (f) Real-ESRGAN, (g) GMM, (h) mVDSR, and (i) mVDSR-GMM

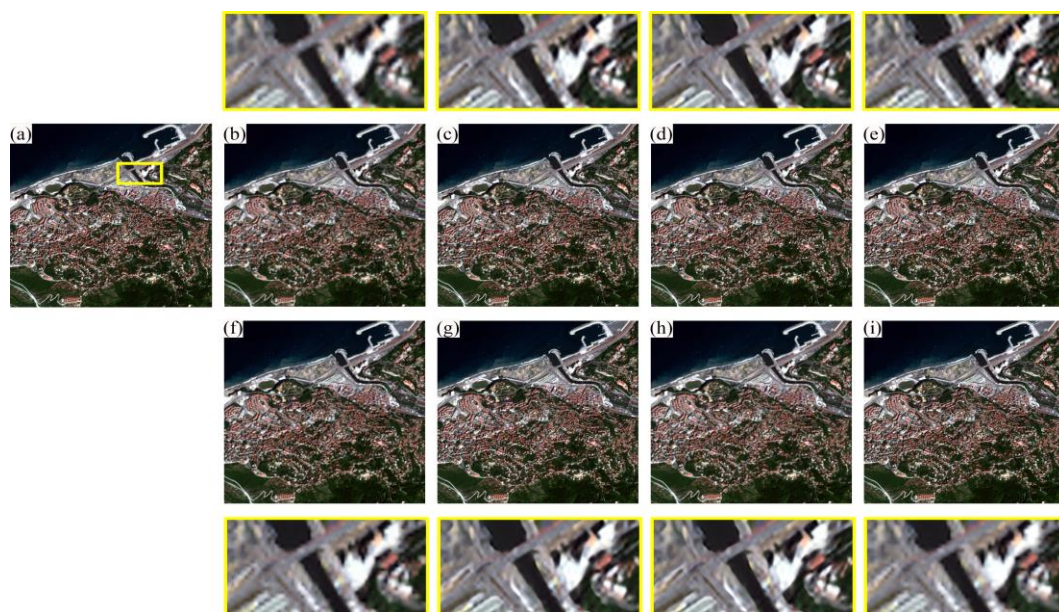


Figure 8. GeoEye-1 x4 scaled results (a) HR image, (b) Bicubic, (c) SwinIR, (d) DSRDCN, (e) SinSR, (f) Real-ESRGAN, (g) GMM, (h) mVDSR, and (i) mVDSR-GMM



Figure 9. QuickBird-2 x2 scaled results (a) HR image, (b) Bicubic, (c) SwinIR, (d) DSRDCN, (e) SinSR, (f) Real-ESRGAN, (g) GMM, (h) mVDSR, and (i) mVDSR-GMM

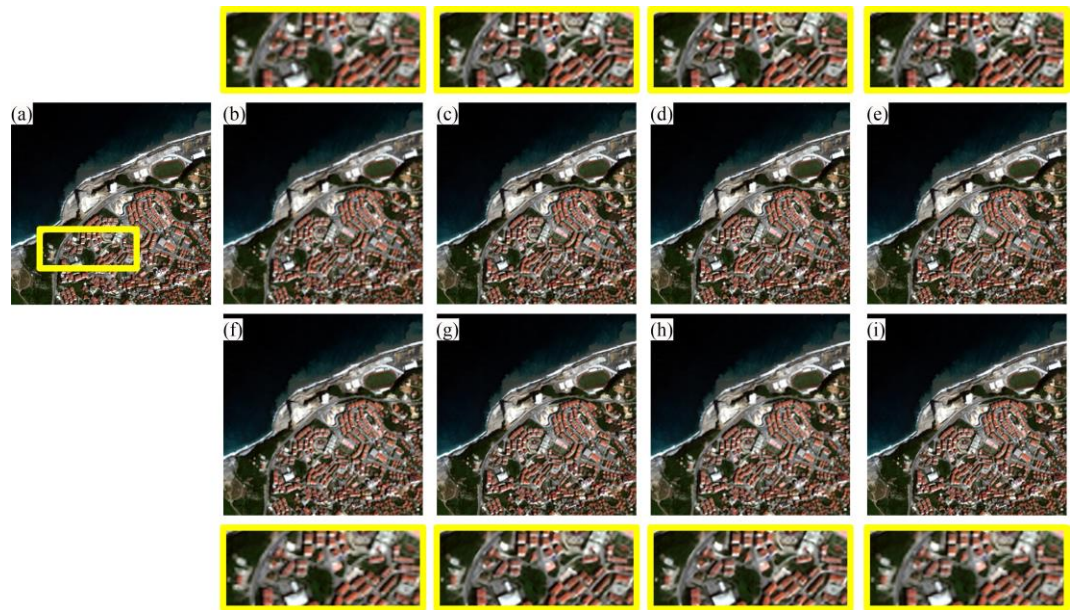


Figure 10. QuickBird-2 x3 scaled results (a) HR image, (b) Bicubic, (c) SwinIR, (d) DSRDCN, (e) SinSR, (f) Real-ESRGAN, (g) GMM, (h) mVDSR, and (i) mVDSR-GMM

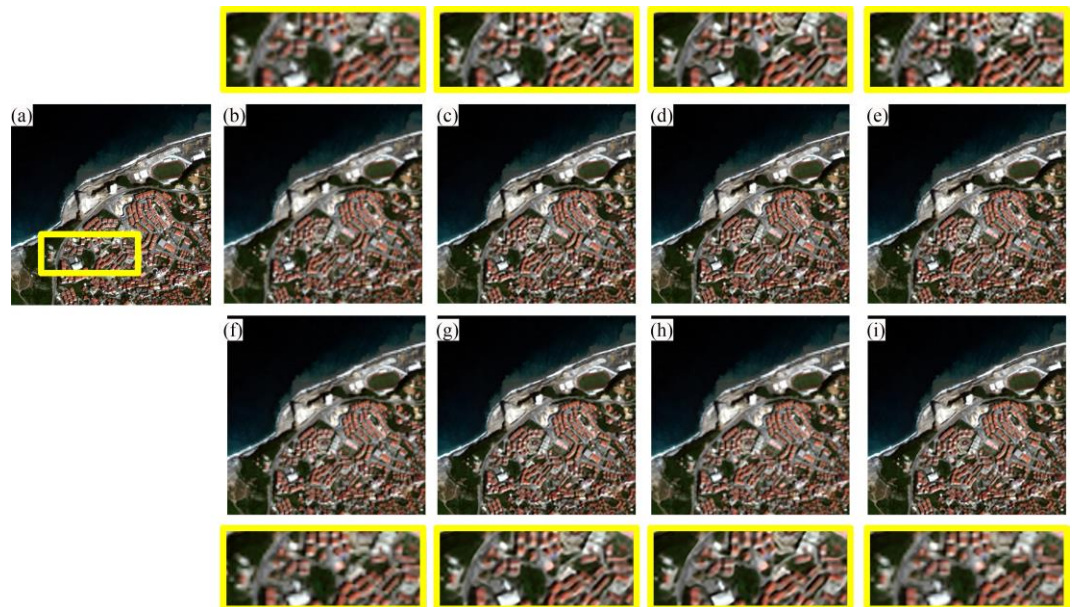


Figure 11. QuickBird-2 x4 scaled results (a) HR image, (b) Bicubic, (c) SwinIR, (d) DSRDCN, (e) SinSR, (f) Real-ESRGAN, (g) GMM, (h) mVDSR, and (i) mVDSR-GMM

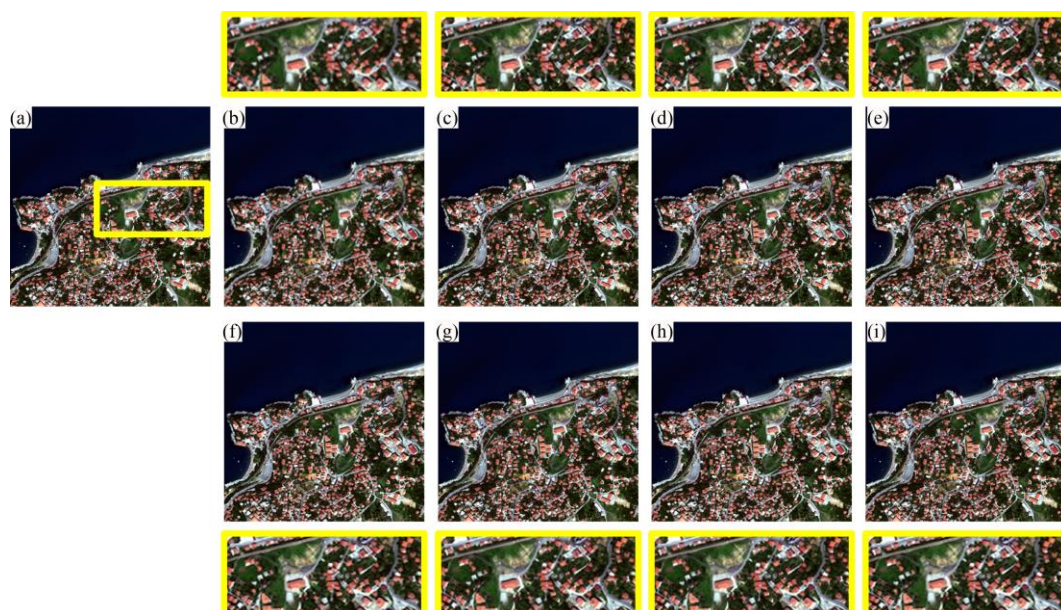


Figure 12. Ikonos-2 x2 scaled results (a) HR image, (b) Bicubic, (c) SwinIR, (d) DSRDCN, (e) SinSR, (f) Real-ESRGAN, (g) GMM, (h) mVDSR, and (i) mVDSR-GMM

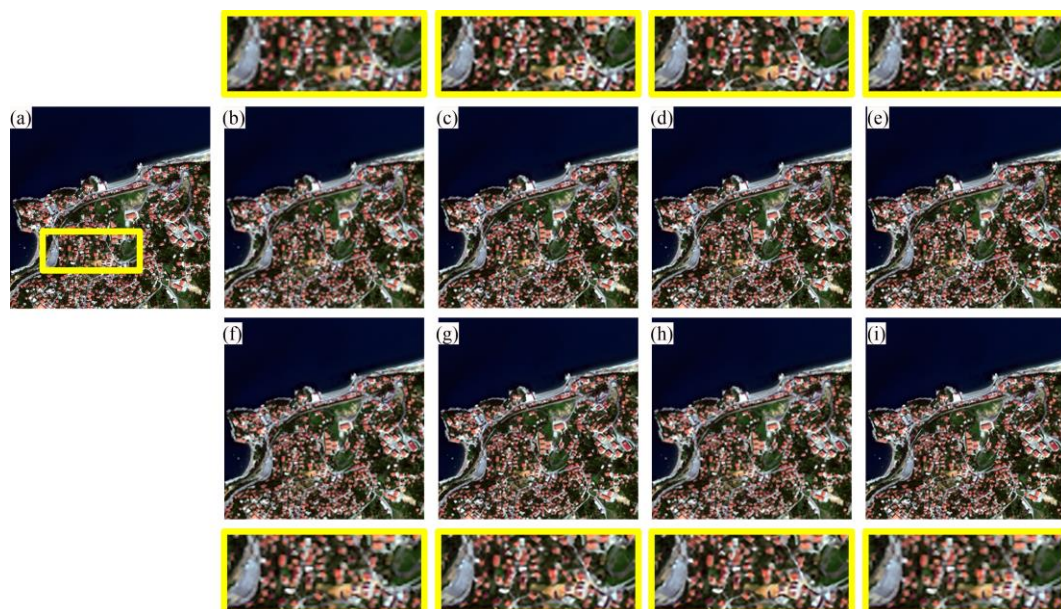


Figure 13. Ikonos-2 x3 scaled results (a) HR image, (b) Bicubic, (c) SwinIR, (d) DSRDCN, (e) SinSR, (f) Real-ESRGAN, (g) GMM, (h) mVDSR, and (i) mVDSR-GMM

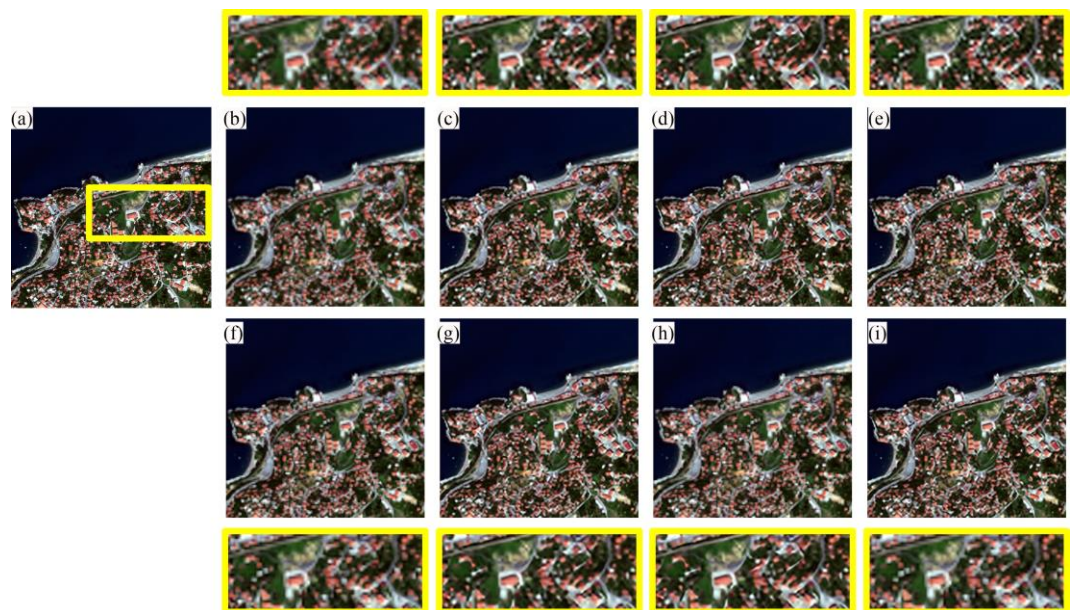


Figure 14. Ikonos-2 x4 scaled results (a) HR image, (b) Bicubic, (c) SwinIR, (d) DSRDCN, (e) SinSR, (f) Real-ESRGAN, (g) GMM, (h) mVDSR, and (i) mVDSR-GMM

Using SR image in land use cover analysis is of significant importance due to its ability to enhance the resolution of satellite imagery. SR of the image improves the level of detail in the images, making it easier to identify and classify specific land features. This is crucial in land use cover analysis, where accurate identification of objects such as buildings, roads, and vegetation is essential. The literature has proposed various methodologies for extracting man-made objects, specifically focusing on road networks and buildings. It is possible to separate them as machine learning-based (Dornaika et al., 2016; Misra et al., 2020), deep learning-based (Ayala et al., 2021; Fan et al., 2021; Sheikh et al., 2022), and segmentation-based (Ünsalan & Boyer, 2005; Ok, 2013; Grinias et al., 2016). Although these methods produce high classification performance, they can be cumbersome in producing fast results. Given that this does not constitute the primary focus of this article, a threshold-based method was used due to its ability to produce rapid and practical results. Threshold methods can be employed to distinguish buildings and roads from sea and green areas, but their success depends on specific conditions and imaging characteristics. To achieve this differentiation, selecting an appropriate threshold value is crucial. This research utilized the histogram-curve fitting threshold method to evaluate the performance of mVDSR-GMM, which yielded optimal results. In (Günen, 2023), the thresholding method based on the fit of the histogram curve is reported to yield the best results compared to nine traditional thresholding methods. In this study, cubic spline was preferred as it provided the best results in the histogram-curve-fitting-based thresholding method. The results obtained for different scale levels of the methods have been super-imposed on the HR image and the results of their difference images are presented in Figure 15. The classification performances of all methods were evaluated with accuracy, specificity, F-measure and recall metrics, as well as visual research. The results are presented in Figure 16. To obtain the relevant metrics, 5000 pixels were randomly sampled from the sea and green areas, and 5000 pixels were randomly sampled from roads and buildings. Although mVDSR-GMM achieves the best overall performance across the evaluation metrics for all scaling factors, the results indicate that all tested super-resolution methods exhibit robust capabilities. This makes it difficult to distinguish a single superior method under all conditions, highlighting the competitive nature of these approaches. The incremental advantages associated with mVDSR-GMM underscore its efficacy, especially in endeavors necessitating precise feature preservation and high-quality reconstruction. It can be seen that the increased spatial resolution provided by image SR contributes to more

accurate classification of land cover types. This is particularly beneficial in distinguishing between land use classes that may have intricate features. In urban areas where land use is diverse and complex, SR helps capture finer details.

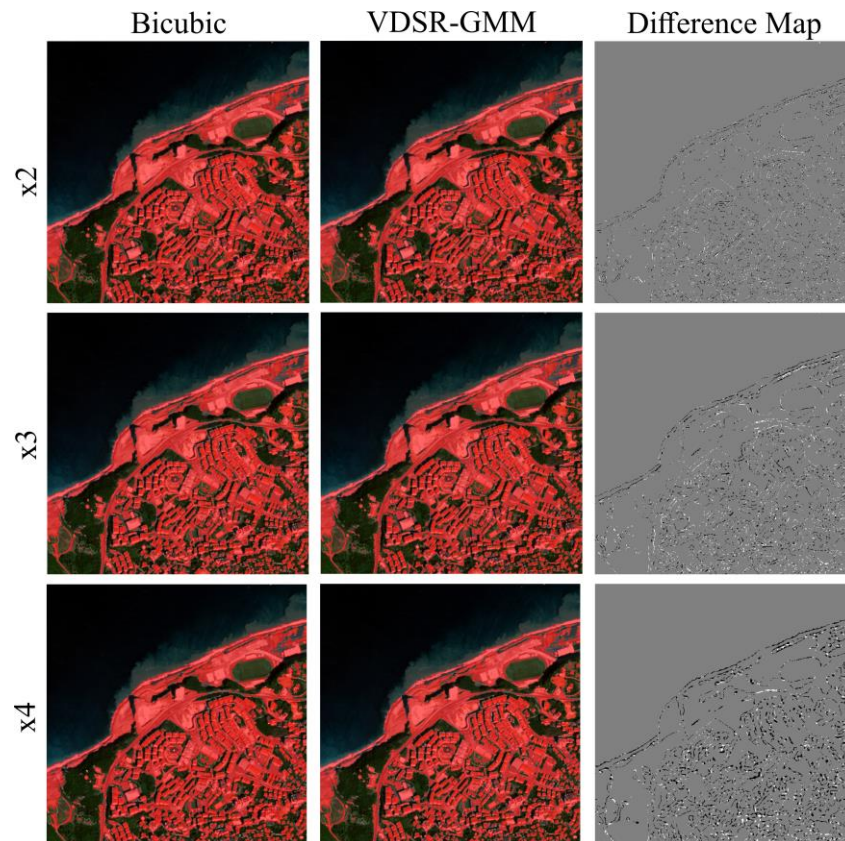


Figure 15. Threshold results of Bicubic and mVDSR-GMM based methods at varying scale

This is vital for urban planning, monitoring infrastructure development, and understanding the dynamics of urban landscapes. In urban areas, man-made structures exhibit high density and regular patterns, whereas in rural regions, it is not uncommon to encounter only a few buildings scattered across vast distances, connected by a sparse network of frequently unpaved roads. The integration of the GMM component into more recent state-of-the-art super-resolution architectures could be systematically investigated. This exploration may enable further performance improvements by leveraging the strengths of both statistical modeling and advanced neural network designs. Additionally, such an approach could provide deeper insights into the compatibility of traditional probabilistic methods with modern deep learning frameworks, thereby expanding the scope and impact of hybrid methodologies in SR tasks.

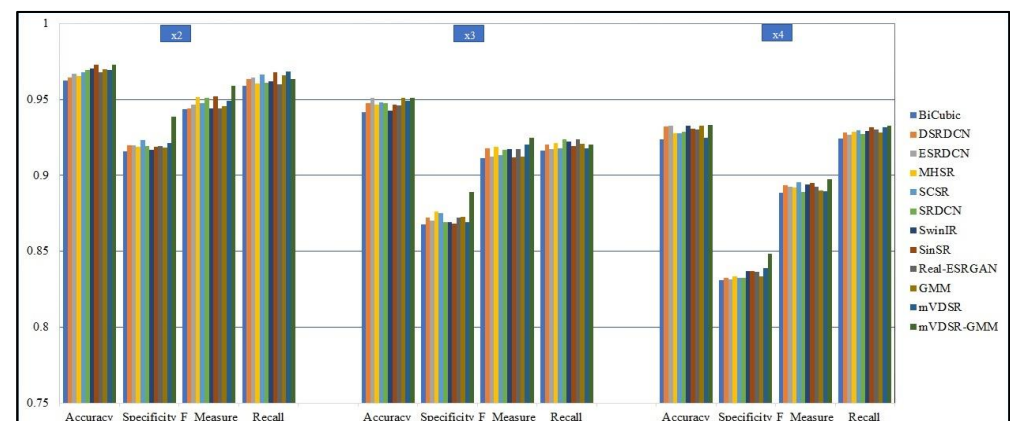


Figure 16. Performance comparison of SISR methods for land use cover analysis

The proposed mVDSR-GMM framework possesses specific structural and computational limitations. The 25-layer architecture requires significant memory and computational resources when processing high-resolution satellite imagery. Furthermore, the performance of the model is sensitive to the characteristics of the training data, and the effectiveness of the GMM component relies on the accurate modeling of image features with Gaussian components; any significant deviation from this distribution may impact reconstruction accuracy. Crucially, the model faces challenges in capturing long-range spatial dependencies. This limitation stems from the local receptive fields inherent in standard convolutional layers and the patch-based nature of GMMs, which constrain the network's ability to integrate global contextual information and correlate distant structural features across the image. However, the adaptive learning capability of GMMs enables the effective capture of complex local patterns, while their inherent noise robustness contributes to the stability of the super-resolution process, particularly in the presence of noisy input data.

4. Conclusion

In this study, the proposed mVDSR-GMM method was compared with Bicubic, DSRDCN, ESRDCN, MHSR, SCSR, SRDCN, SwinIR, SinSR, Real-ESRGAN, GMM, and VDSR methods to assess its effectiveness in improving the spatial resolution of satellite imagery. mVDSR training was carried out on diverse data sets, including DIV2K and aerial photographs from the Turkish General Directorate of Mapping. The performance of the methods has been evaluated through experiments on diverse satellite image datasets, including GeoEye-1, QuickBird-2, and Ikonos-2, which have been used for extensive experiments and evaluations to assess the generalizability and robustness of the proposed method. Although mVDSR-GMM generally outperformed the other techniques, obtaining lower *Brisque*, *SAM*, *FD*, and *MSE* as well as higher *PSNR*, *UIQI*, and *SSIM* statistical scores, the least effective method was Bicubic. In particular, this applies to the $\times 2$ and $\times 3$ scale factors. However, approaches such as SwinIR and SinSR operate well in some contexts, particularly where visual clarity and subjective quality are prioritized. The mVDSR-GMM framework achieved a measurable increase in reconstruction quality, as objectively validated by the quantitative improvements in *PSNR*, *SSIM*, and *FD* metrics across all evaluated satellite datasets. From a remote sensing perspective, the implications of this work are significant. The mVDSR-GMM framework provides a robust, software-based solution to the fundamental hardware trade-off between spatial and spectral resolution in Earth observation satellites. By improving image clarity and edge sharpness, the proposed method directly enhances the accuracy of land use/cover analysis, specifically in the binary separation of urban structures (buildings and roads) from natural environments. This discrimination was facilitated by applying the histogram-curve-fitting threshold method. The proposed methodologies may be suitably applied to contexts in which images are utilized for human interpretative tasks, such as cartographic mapping and visual representation. Further studies could aim to improve the model design, investigate a broader range of datasets, and address particular issues in practical applications.

The effects of applying the proposed techniques to Sentinel-2 and Landsat-9 satellite data, which have comparatively lower spatial resolutions, will be examined. Furthermore, upcoming research will consider the incorporation of GMM into deep learning architectures beyond VDSR to enhance its utility and effectiveness.

Author Contributions: The author confirms sole responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation.

AI Contribution Statement: No artificial intelligence (AI) tools were used in the preparation of this manuscript. All work was conducted entirely by the author.

Funding: This study received no external funding.

Research and publication ethics statement: In the study, the author declares that there is no violation of research and publication ethics and that the study does not require ethics committee approval.

Conflicts of Interest: The author declares no conflicts of interest.

Data Availability: The data presented in this study are available on request from the corresponding author.

References

- Agustsson, E., & Timofte, R. (2017). Ntire 2017 challenge on single image super-resolution: Dataset and study. *Proceedings Book of the IEEE conference on computer vision and pattern recognition workshops*, Honolulu, Hawaii, USA. 126-135.
- Al-Mekhlafi, H., & Liu, S. (2024). Single image super-resolution: A comprehensive review and recent insight. *Frontiers of Computer Science*, 18, 181702. <https://doi.org/10.1007/s11704-023-2588-9>
- Arabi Aliabad, F., Ghafarian Malmiri, H., Sarsangi, A., Sekertekin, A., & Ghaderpour, E. (2023). Identifying and monitoring gardens in urban areas using aerial and satellite imagery. *Remote Sensing*, 15(16), 4053. <https://doi.org/10.3390/rs15164053>
- Atasever, U. H., & Günen, M. A. (2021). Change detection approach for SAR imagery based on arc-tangential difference image and k-Means++. *IEEE Geoscience and Remote Sensing Letters*, 19, 1–5. <https://doi.org/10.1109/LGRS.2021.3127964>
- Ayala, C., Sesma, R., Aranda, C., & Galar, M. (2021). A deep learning approach to an enhanced building footprint and road detection in high-resolution satellite imagery. *Remote Sensing*, 13(16), 3135. <https://doi.org/10.3390/rs13163135>
- Chen, C., & Fowler, J. E. (2012). Single-image super-resolution using multihypothesis prediction. *Forty Sixth Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, 608–612.
- Chen, Y.-K., Cheng, F.-C., & Tsai, P. (2011). A gray-level clustering reduction algorithm with the least PSNR. *Expert Systems with Applications*, 38(8), 10183–10187. <https://doi.org/10.1016/j.eswa.2011.02.071>
- Chen, Y., Liu, L., Phonevilay, V., Gu, K., Xia, R., Xie, J., Zhang, Q., & Yang, K. (2021). Image super-resolution reconstruction based on feature map attention mechanism. *Applied Intelligence*, 51, 4367–4380. <https://doi.org/10.1007/s10489-020-02116-1>
- Cheng, M., Ma, H., Ma, Q., Sun, X., Li, W., Zhang, Z., Sheng, X., Zhao, S., Li, J., & Zhang, L. (2023). Hybrid transformer and CNN attention network for stereo image super-resolution. *Proceedings Book of IEEE Conference on Computer Vision and Pattern Recognition*, Vancouver, Canada, 1702–1711.
- Civicioglu, P., & Besdok, E. (2022). Contrast stretching based pansharpening by using weighted differential evolution algorithm. *Expert Systems with Applications*, 208, 118144. <https://doi.org/10.1016/j.eswa.2022.118144>
- Civicioglu, P., & Besdok, E. (2024). Pansharpening of remote sensing images using dominant pixels. *Expert Systems with Applications*, 242, 122783. <https://doi.org/10.1016/j.eswa.2023.122783>
- Courtrai, L., Pham, M.-T., & Lefèvre, S. (2020). Small object detection in remote sensing images based on super-resolution with auxiliary generative adversarial networks. *Remote Sensing*, 12(19), 3152. <https://doi.org/10.3390/rs12193152>
- Dabbech, A., Terris, M., Jackson, A., Ramatsoku, M., Smirnov, O. M., & Wiaux, Y. (2022). First AI for deep super-resolution wide-field imaging in radio astronomy: Unveiling structure in ESO 137-006. *The Astrophysical Journal Letters*, 939, L4. <https://doi.org/10.3847/2041-8213/ac98af>
- Dai, Q., Cheng, X., Qiao, Y., & Zhang, Y. (2020). Crop leaf disease image super-resolution and identification with dual attention and topology fusion generative adversarial network. *IEEE Access*, 8, 55724–55735. <https://doi.org/10.1109/ACCESS.2020.2982055>
- Dong, C., Loy, C. C., He, K., & Tang, X. (2014). Learning a deep convolutional network for image super-resolution. *European Conference on Computer Vision*. Springer.
- Dornaika, F., Moujahid, A., El Merabet, Y., & Ruichek, Y. (2016). Building detection from orthophotos using a machine learning approach: An empirical study on image segmentation and descriptors. *Expert Systems with Applications*, 58, 130–142. <https://doi.org/10.1016/j.eswa.2016.03.024>
- Fan, Y., Ding, X., Wu, J., Ge, J., & Li, Y. (2021). High spatial-resolution classification of urban surfaces using a deep learning method. *Building and Environment*, 200, 107949. <https://doi.org/10.1016/j.buildenv.2021.107949>
- Farooq, M., Dailey, M. N., Mahmood, A., Moonrinta, J., & Ekpanyapong, M. (2021). Human face super-resolution on poor quality surveillance video footage. *Neural Computing and Applications*, 33, 13505–13523. <https://doi.org/10.1007/s00521-021-05973-0>
- Glasner, D., Bagon, S., & Irani, M. (2009). Super-resolution from a single image. *IEEE Conference on Computer Vision and Pattern Recognition*. Springer.
- Grinias, I., Panagiotakis, C., & Tziritas, G. (2016). MRF-based segmentation and unsupervised classification for building and road detection in peri-urban areas of high-resolution satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 122, 145–166. <https://doi.org/10.1016/j.isprsjprs.2016.10.010>
- Günen, M. A. (2021). Weighted differential evolution algorithm based pansharpening. *International Journal of Remote Sensing*, 42(21), 8468–8491. <https://doi.org/10.1080/01431161.2021.1976874>

- Günen, M. A. (2022). Performance comparison of deep learning and machine learning methods in determining wetland water areas using EuroSAT dataset. *Environmental Science and Pollution Research*, 29, 21092–21106. <https://doi.org/10.1007/s11356-021-17177-z>
- Günen, M. A. (2023). Comparison of histogram-curve fitting-based and global threshold methods for cloud detection. *International Journal of Environmental Science and Technology*, 1–26. <https://doi.org/10.1007/s13762-023-05379-6>
- Günen, M. A., Pérez-Delgado, M.-L., & Beşdok, E. (2024). L0-Norm based image pansharpening by using population-based algorithms. *AIMS Mathematics*, 9(1), 32578–32628. <https://doi.org/10.3934/math.20241561>
- He, J., Yuan, Q., Li, J., Xiao, Y., Liu, D., Shen, H., & Zhang, L. (2023). Spectral super-resolution meets deep learning: Achievements and challenges. *Information Fusion*, 101, 101812. <https://doi.org/10.1016/j.inffus.2023.101812>
- Jiang, J., Wang, C., Liu, X., & Ma, J. (2021). Deep learning-based face super-resolution: A survey. *ACM Computing Surveys*, 55(1), 1–36. <https://doi.org/10.1145/3485132>
- Joshi, O., Joshi, A., & Sawant, S. (2022). 2D MRI super resolution using generative adversarial network. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4291804>
- Kim, J., Lee, J. K., & Lee, K. M. (2016). Accurate image super-resolution using very deep convolutional networks. *Proceedings Book of the IEEE conference on computer vision and pattern recognition*, Las Vegas, Nevada, USA, 1646–1654.
- Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., & Shi, W. (2017). Photo-realistic single image super-resolution using a generative adversarial network. *Proceedings Book of the IEEE conference on computer vision and pattern recognition*, Honolulu, Hawaii, USA, 4681–4690.
- Lei, S., Shi, Z., & Mo, W. (2021). Transformer-based multistage enhancement for remote sensing image super-resolution. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–11. <https://doi.org/10.1109/TGRS.2021.3136190>
- Leinonen, J., Nerini, D., & Berne, A. (2020). Stochastic super-resolution for downscaling time-evolving atmospheric fields with a generative adversarial network. *IEEE Transactions on Geoscience and Remote Sensing*, 59(9), 7211–7223. <https://doi.org/10.1109/TGRS.2020.3032790>
- Lepcha, D. C., Goyal, B., Dogra, A., & Goyal, V. (2023). Image super-resolution: A comprehensive review, recent trends, challenges and applications. *Information Fusion*, 91, 230–260. <https://doi.org/10.1016/j.inffus.2022.10.007>
- Misra, M., Kumar, D., & Shekhar, S. (2020). Assessing machine learning based supervised classifiers for built-up impervious surface area extraction from Sentinel-2 images. *Urban Forestry & Urban Greening*, 53, 126714. <https://doi.org/10.1016/j.ufug.2020.126714>
- Mittal, A., Moorthy, A. K., & Bovik, A. C. (2011). Blind/referenceless image spatial quality evaluator. *Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, 723–727. <https://doi.org/10.1109/ACSSC.2011.6190099>
- Ogawa, Y., Ariki, Y., & Takiguchi, T. (2012a). Super-resolution by GMM based conversion using self-reduction image. *IEEE International Conference on Acoustics, Speech and Signal Processing*, Kyoto, Japan, 2012, 1285–1288. <https://doi.org/10.1109/ICASSP.2012.6288124>
- Ogawa, Y., Hori, T., Takiguchi, T., & Ariki, Y. (2012b). Super-resolution using GMM and PLS regression. *IEEE International Symposium on Multimedia*, Irvine, CA, USA, 298–301. <https://doi.org/10.1109/ISM.2012.62>
- Ok, A. O. (2013). Automated detection of buildings from single VHR multispectral images using shadow information and graph cuts. *ISPRS Journal of Photogrammetry and Remote Sensing*, 86, 21–40. <https://doi.org/10.1016/j.isprsjprs.2013.09.004>
- Özyurt, F., Sert, E., & Avcı, D. (2020). An expert system for brain tumor detection: Fuzzy C-means with super resolution and convolutional neural network with extreme learning machine. *Medical Hypotheses*, 134, 109433. <https://doi.org/10.1016/j.mehy.2019.109433>
- Pérez-Delgado, M.-L., & Günen, M. A. (2023). A comparative study of evolutionary computation and swarm-based methods applied to color quantization. *Expert Systems with Applications*, 224, 120666. <https://doi.org/10.1016/j.eswa.2023.120666>
- Ran, Q., Xu, X., Zhao, S., Li, W., & Du, Q. (2020). Remote sensing images super-resolution with deep convolution networks. *Multimedia Tools and Applications*, 79, 8985–9001. <https://doi.org/10.1007/s11042-018-7091-1>
- Salguero-Cruz, A., Latorre-Carmona, P., & García-Osorio, C. I. (2022). Proposal of a comparative framework for face super-resolution algorithms in forensics. *Iberian Conference on Pattern Recognition and Image Analysis*. Springer.
- Sandeep, P., & Jacob, T. (2016). Single image super-resolution using a joint GMM method. *IEEE Transactions on Image Processing*, 25(9), 4233–4244. <https://doi.org/10.1109/TIP.2016.2588319>
- Sheikh, M. A. A., Maity, T., & Kole, A. (2022). Man-made object extraction from remote sensing images using Gabor energy features and probabilistic neural networks. *ICTACT Journal on Image & Video Processing*, 13(1). <https://doi.org/10.21917/ijivp.2022.0407>
- Singh, A., & Singh, J. (2020). Survey on single image based super-resolution—implementation challenges and solutions. *Multimedia Tools and Applications*, 79, 1641–16472. <https://doi.org/10.1007/s11042-019-08254-0>
- Şan, M., Akpınar, A., Bingölbali, B., & Kankal, M. (2021). Geo-spatial multi-criteria evaluation of wave energy exploitation in a semi-enclosed sea. *Energy*, 214, 118997. <https://doi.org/10.1016/j.energy.2020.118997>

- Tai, Y., Yang, J., & Liu, X. (2017). Image super-resolution via deep recursive residual network. *Proceedings Book of the IEEE conference on computer vision and pattern recognition workshops*, Honolulu, Hawaii, USA, 3147–3155.
- Ünsalan, C., & Boyer, K. L. (2005). A system to detect houses and residential street networks in multispectral satellite images. *Computer Vision and Image Understanding*, 98(3), 423–461. <https://doi.org/10.1016/j.cviu.2004.10.006>
- Wang, P., Bayram, B., & Sertel, E. (2022a). A comprehensive review on deep learning based remote sensing image super-resolution methods. *Earth-Science Reviews*, 232, 104110. <https://doi.org/10.1016/j.earscirev.2022.104110>
- Wang, Y., Bashir, S. M. A., Khan, M., Ullah, Q., Wang, R., Song, Y., Guo, Z., & Niu, Y. (2022b). Remote sensing image super-resolution and object detection: Benchmark and state of the art. *Expert Systems with Applications*, 197, 116793. <https://doi.org/10.1016/j.eswa.2022.116793>
- Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4), 600–612. <https://doi.org/10.1109/TIP.2003.819861>
- Yang, J., Wright, J., Huang, T. S., & Ma, Y. (2010). Image super-resolution via sparse representation. *IEEE Transactions on Image Processing*, 19(11), 2861–2873. <https://doi.org/10.1109/TIP.2010.2050625>
- Zhang, Y., Tian, Y., Kong, Y., Zhong, B., & Fu, Y. (2018). Residual dense network for image super-resolution. *Proceedings Book of the IEEE conference on computer vision and pattern recognition*, Salt Lake City, Utah, USA, 2472–2481.