Zakia Kezzoula
Djamel Gaceb

PSO-WESRGAN: A NOVEL DOCUMENT IMAGE SUPER RESOLUTION

Abstract *One of the major challenges of document images that can hinder readability and the analysis of information is low resolution; this is typically caused by low-pixel density scanning or excessive compression to save storage space. This results in a loss of fine detail in images, making it difficult to detect critical information. To solve these problems, super-resolution techniques are used. These techniques improve image quality by increasing the resolution while maintaining the fine detail. PSO-WESRGAN is an innovative method that combines wavelet processing, deep-transfer learning, and particle swarm optimization (PSO). Wavelet processing analyzes image detail at diverse scales and orientations, while transfer-based deep-learning advantages pre-trained models on vast image data sets. By integrating PSO, the efficiency of the method is enhanced through the optimal exploration of the solution space to identify the best parameters for the super-resolution model. The experimental results show the effectiveness of this method and open up prospects for future improvements in the super-resolution of document images.*

Keywords document image, super-resolution, transfer deep learning, particle swarm optimization, wavelet transform

Citation Computer Science 26(4) 2025: 5–31

1. Introduction

Digitized documents represent a valuable source of information in many sectors; their high resolution is essential for making text legible and illustrations visible, thus facilitating subsequent accurate automatic analysis. However, their quality can be compromised by various forms of degradation; among these, the low resolution that results from technical limitations during digitization compromises the accuracy of fine details. Figure 1 shows low-resolution image examples.

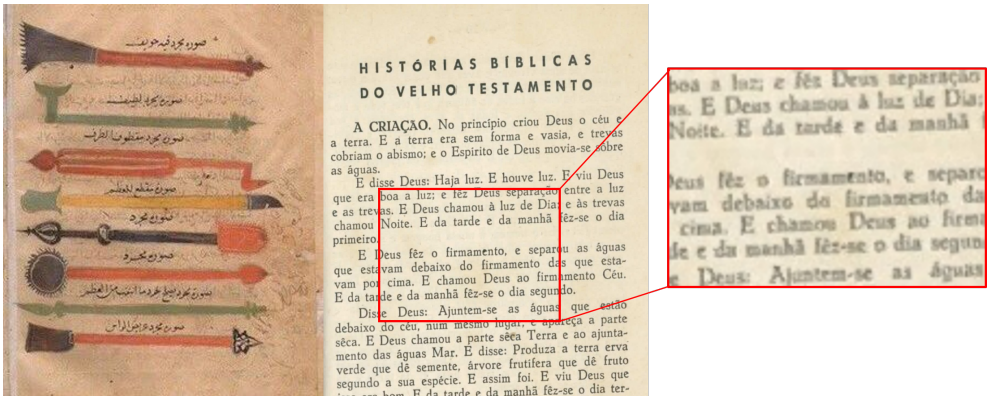

Figure 1. Low-resolution image examples

This requires advanced image-processing techniques to restore the visual quality of documents, improve their legibility, and facilitate the accurate extraction of the information they contain. This article focuses on the development of a super-resolution approach using bio-inspired and deep pre-processing to improve document image quality and facilitate the extraction of useful information. Super-resolution aims to increase the resolution of images to reveal fine details and improve their legibility. The development of image super-resolution methods has distinct changes in levels through the phases and the corresponding technological changes. These techniques are classified into two broad families: Single-Image Super-Resolution (SISR), and Multi-Image Super-Resolution (MISR). The MISR methods increase image resolution by using a series of LR image scenes. These methods take advantage of information redundancy across multiple images to produce a high-resolution (HR) image. Although multi-image super-resolution methods may offer significant improvements in high-resolution image quality in various application fields such as medical [5, 8, 24, 42] and satellite [3, 33, 38, 41], they face significant challenges related to computational complexity, data availability, and motion-artifact management. These aspects should be carefully assessed when choosing a super-resolution method for a given scenario. The second family generates a high-resolution output image using a single low-resolution input image. These methods offer considerable potential for improving image quality and resolution – especially for super-resolution document images. In this category,

we can identify two subcategories of techniques: interpolation-based techniques, and learning-based techniques (using machine or deep-learning methods). Using mathematical techniques, interpolation techniques determine the values of missing pixels in a LR picture. Although interpolation methods are simple, they can cause noise, blurring, and false colors and may not capture fine details in high-resolution images (especially thin lines of text). In our study, we compared our proposed approach with the following interpolation methods: Bell [28], Bi-linear [36], Bspline [13], Contour stencils [11,12], Gaussien [34], Hanning [14], Lanczos [10], neighbor [1], Mitchell [26], and bicubic [15]. Learning-based techniques use machine-learning models' abilities to convert low-resolution (LR) images into high-resolution (HR) images, thus surpassing traditional interpolation techniques in terms of quality and accuracy.

Documents often contain fine text, diagrams, and details that are essential for their legibility and accurate interpretation. Low-resolution images frequently suffer from noise, artifacts, and losses of detail; these can cause errors during analysis or optical character recognition (OCR). In addition, preserving the original layout is crucial for applications such as archiving or document reuse. The super-resolution of document images is a major challenge due to the unique requirements of this type of image. The difficulty is in the ability to reconstruct these textual and visual details while maintaining the integrity of the document structure. Our method effectively addresses the challenges of the super-resolution of document images by combining wavelet decomposition, pre-driven ESRGAN models, and PSO optimization. Wavelet decomposition helps preserve critical details, while ESRGAN networks improve visual quality by reconstructing complex textures. PSO optimization ensures an optimal balance between different sub-bands for a faithful reconstruction. This approach also improves OCR performance by maximizing readability and preserving the original layout, providing a robust solution for demanding document applications. The next parts of the article will review existing deep-learning-based super-resolution techniques and provide a detailed description of our approach to super-resolution document images. In addition, we will discuss the experimental results that support the effectiveness of our approach to improve document image resolution.

2. Deep-learning-based super-resolution techniques

Super-resolution models using deep learning have demonstrated exceptional performance in a number of quality metrics. The most notable approaches in terms of quality and complexity are: Super-Resolution Convolutions Neural Network (also known as SRCNN) was developed in [7]. SRCNN is a supervised-learning technique that allows learning the non-linear correspondence between low-resolution and high-resolution images. The main layers that define the architecture of the SRCNN model are the convolution, pooling, and reconstruction layers. The authors of SRCNN demonstrated that it gives superior results by enhancing the details and quality of the generated images. Improved Deep Super-Resolution (IDSR) (discussed in [22]) is characterized by its rather profound neural network architecture proposed to solve

the degradation issue due to depth by using a solution known as residual connections. Using residual blocks, Very Deep Super-Resolution (VDSR) [6] is a highly accurate super-resolution approach using a deep convolution network; this was inspired by the VGG network for ImageNet classification. Composed of 20 layers, this model efficiently exploits contextual information over large image regions in a cascade of small filters. To speed up training, high learning rates and gradient clipping are used, ensuring training stability. The method has shown superior performance in terms of accuracy and visual improvements over existing methods.

In [48], the authors described the Residual Dense Network for image super-resolution. This method attempts to employ dense residual blocks to enhance super-resolution (SR) image quality-learning residual details. Depending on the density of these residuals, the proposed approach envisages enhancing the details of HR images using dense residual blocks. RCAN (Residual Channel Attention Networks) is a type of SR-based deep learning [47]. Some of the unique features of the chosen RCAN architecture concern the usage of residual blocks and channel attention. RCAN aims to extract the essential features of the problem while maintaining a compact model. The fundamental units of RCAN are deep residual blocks. Each residual block contains multiple convolution layers, which helps in learning highly non-linear features. The RCAN method introduces the channel-attention mechanism; this mechanism lets the network focus on the most informative channels (or features) in each residual block. It facilitates the identification of important channels for super-resolution and the efficient allocation of computational resources. Fast and Accurate Image Super-Resolution with Deep Laplacian Pyramid Networks [19] exploits information at different scales to reconstruct detailed super-resolved images. Deep Plug-and-Play Super-Resolution for Arbitrary Blur Kernels [46] uses neural networks to estimate and invert arbitrary blurs, thus improving image resolution. Uformer: A General U-Shaped Transformer for Image Restoration [45] proposes a novel method for image-restoration tasks that makes use of the Uformer, which is an adaptable U-shaped transformer model specifically designed for image-restoration tasks. Uformer is a model that includes an encoder-decoder structure with skip links. The encoder component extracts multi-scale information from an input image, while the decoder component reconstructs the image. The various scales can be reconstructed through skip connections to create the image since they help the model maintain and integrate information from different scales. Super-resolution on a single image using a directional variance attention network [2] proposes the directional variance-based attention network (or DiVANet) for image super-resolution. DiVANet presents a new strategy for SR that includes an optimized attention mechanism (referred to as DiVA) and another system known as the Residual Attention Feature Group (RAFG). This enhances the quality of super-resolution, as it works with spatial features and ensures that fine details are well-handled; all of this is achieved with the model being light, fast, and efficient. Single super-resolution images by the progressive integration of orientation-sensitive features [16] provides an SISR-PF-OA that uses an OAM feature-extraction technique using both 1D and 2D convolution kernels

with channel attention. It also presents progressive fusion to enhance multi-scale fusion performance, making it potential to offer high-quality output with high accuracy and less computational complexity as compared to the other deep-learning methods. SR-FEINR: End-to-End Video Super-Resolution Through Feature-Enhanced Implicit Neural Representation [23] provides a new solution for the successive SR of remote-sensing images. The technique consists of three primary modules: an encoding module, a feature-extraction module, and a feature-enhanced multi-layer perceptron. The authors in [21] proposed a method for image restoration known as SwinIR, which was built on Swin Transformer. SwinIR has three main components: surface-feature extractors, deep-feature extractors, and HR-reconstruction modules. The model uses a series of residual Swin Transformer (RSTB) blocks to extract deep characteristics. Each RSTB is designed with a residual connection, convolution layer, and Swin Transformer layer set to improve the extraction of deep features in the restoration process. The authors presented a new transformer-based technique for the super-resolution of images in [4]. They provided a hybrid attention transformer (HAT) that incorporates both window-based self-attention models and channel attention to address geographic information. Multi-Level Dispersion Residual Network for Efficient Image Super-Resolution [25] was designed to improve the efficiency of image super-resolution (EISR). This model builds on EADB, which has ECCA and MDSA added to it. The experiments proved the effectiveness of these techniques. This approach achieved first place in the NTIRE 2023 Efficient SR Challenge. The Super-resolution Generative Adversarial Network (SRGAN) is a super-resolution method that was introduced in [20]. It is based on the GAN architecture, which involves two neural networks, a generator, and a discriminator, which train concurrently. The generator creates a high-resolution version from an input's low-resolution image. The discriminator looks for distinctions between images generated by the generator and actual HR images. The capacity of the generator to really create super-resolution images is enhanced by this adversarial training process.

Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN) [44] are composed of two opposing neural networks: discriminating, and generating neural networks. The generator is responsible for transforming an LR image into a more accurate and detailed HR version. It analyzes large sets of high-resolution images to identify the relationships between the pixels in an image and learns how to reconstruct them accurately. By including dense residual blocks and increasing perceptual loss, the ERGAN model outperformed the SRGAN model.

The discriminator acts as a binary classifier to distinguish real images from generated ones. Designed to operate efficiently in an image-resolution context, the discriminator is built with convolution and pooling layers that allow it to extract discriminant characteristics from input images. Its goal is to maximize its ability to identify subtle differences between real HR images and those generated by the generator.

During the training, the discriminator and generator are adversely optimized: the discriminator adjusts its weights to minimize its error between real and generated

images. The generator tries to produce increasingly realistic images to fool the discriminator. This dynamic leads to a gradual refinement of the skills of both networks until a balance is reached, where the generator produces high-resolution images that are impossible to distinguish from real ones. This process allows for the continuous improvement of image-generated quality.

With its more advanced architecture, ESRGAN is better than the SRGAN model at capturing fine features and producing high-quality images in super-resolution. The ESRGAN model uses dense connections within the RRDB block to improve overall system performance when generating high-resolution images. ESRGAN has proven to produce better results than existing super-resolution techniques. It can produce high-resolution images that look more realistic and natural by capturing details. Bi-ESRGAN: A New Approach for Document Image Super-Resolution Based on Dual Deep Transfer Learning [18] is a new approach to the super-resolution images of low-resolution documents. This method uses two ESRGAN networks on separate image cards (the original image, and the contours of the image using the Laplacian filter); it preserves the texture and graphic properties of document images. The experimental results demonstrate the superiority of "Bi-ESRGAN" over several other approaches on a variety of document images, demonstrating significant progress in super-resolution of document images. Recent advances in deep-learning-based super-resolution models, which surpass traditional methods in terms of quality and accuracy, have led to the active exploration of these methods. Our approach is an improvement of the ESRGAN method by integrating information from wavelet transformations and using the PSO algorithm. This combination allows us to find the best data balance for optimizing the resolutions of document images.

3. Proposed approach

Our approach is a super-resolution technique for images of documents that use deep-learning and bio-inspired methods. Figure 2 demonstrates our technique's architecture and procedures.

3.1. Wavelet transfer

Wavelet transformation divides an image into several sub-bands; the LL (Low-Low) sub-band contains the low frequencies of the image. It represents the overall detail and structure of the image. In general, this is the component that retains the most information from the original image. The LH (Low-High) sub-band combines low frequencies in the horizontal direction and high frequencies in the vertical direction, capturing the horizontal details of the image. In contrast, the HL (High-Low) sub-band combines horizontal high frequencies with vertical low frequencies; it represents vertical image detail. The HH (High-High) sub-band contains horizontal and vertical high frequencies; it represents fine image details such as contours and textures [27,30]. Figure 3 illustrates a wavelet transformation example.

In our method, we chose a decomposition level of 1; this means that the image was decomposed once, thus generating an approximation sub-band and three detailed sub-bands. This choice of decomposition was a strategic compromise between the amount of information retained in the image and computational efficiency. By reducing the complexity of the decomposition to a single level, we were able to maintain the most important details while optimizing the resources needed for processing.

By using wavelets for document images, we separated the fine details from the coarse components, thus improving the readability of the characters and small structures; this facilitated the processing of multi-scale information, reduced noise, and optimized image reconstruction. By independently processing the different frequencies, wavelets helped to preserve details while minimizing artifacts, resulting in a high-quality super-resolution image and better preservation of the document structure.

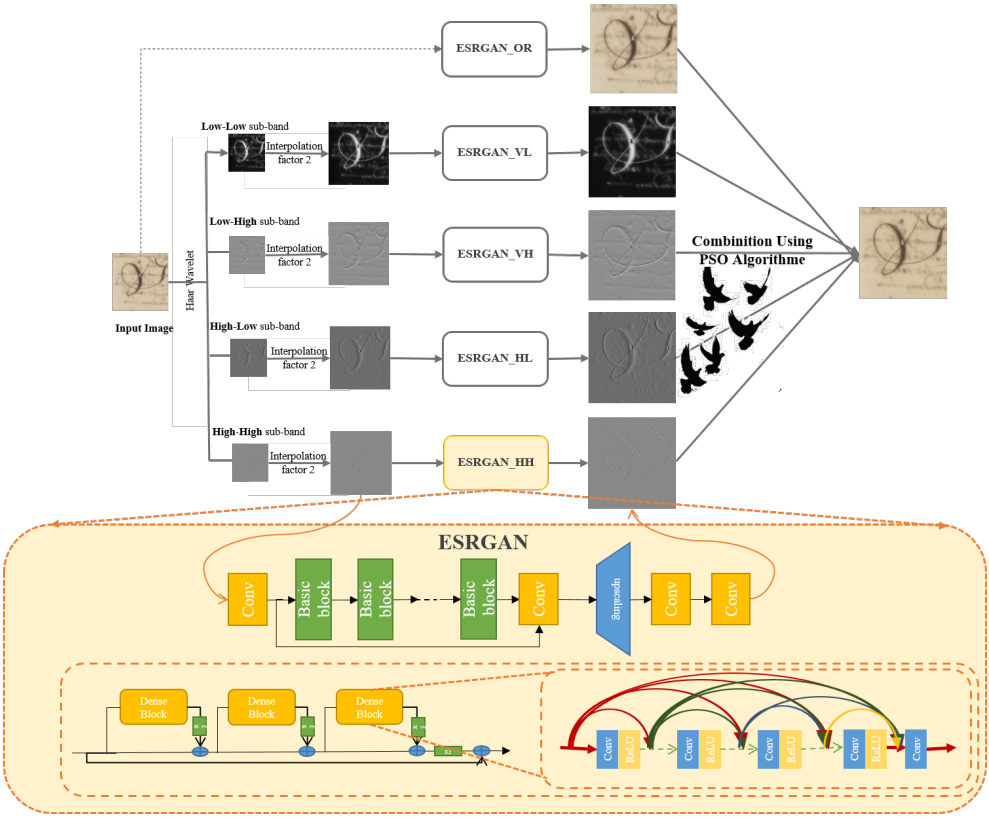

Figure 2. Processes and architecture of our SR technique for document images

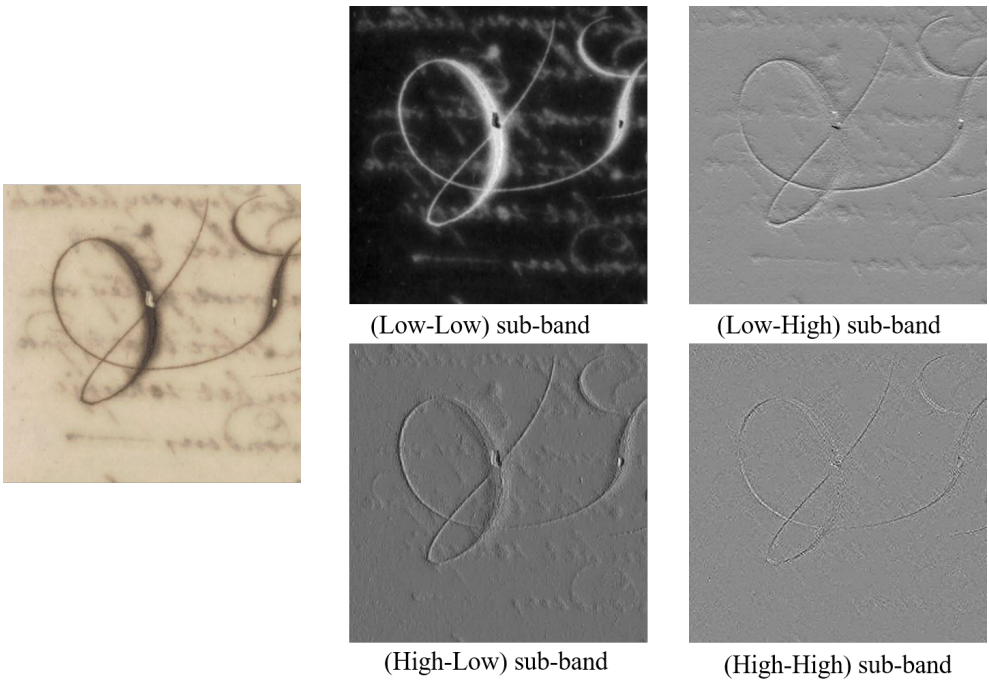

Figure 3. Wavelet transformation example

3.2. Training of ESRGAN networks

We used the ESRGAN model pre-trained on nature images (DIV2K, Flickr2K, and ImageNet) and trained it on the document image database SR_VISIONṪhis database was designed at the LIMOSE laboratory of the University of Boumerdes (UMBB) in order to meet the specific needs of research on super-resolution applied to document images. It provided a practical framework for the training and evaluation of various approaches, taking into account the particularities of documentary images such as the presence of texts, diagrams, and fine details.

The data set was carefully structured into two subsets: TRAIN, and TEST – each consisted of 757 separate images. This organization ensured a robust learning process and an independent evaluation of the model's performance on unseen data.

The images included in this database covered a variety of structures and qualities, thus providing a realistic challenge for super-resolution models and contributing to reliable assessments in practical scenarios.

Figure 4 shows an illustration from the SR_VISION data set.

The SR_VISION data set is not public, and access to it is restricted. Due to confidentiality considerations and internal policy, this data set is for internal use only and cannot be shared with external parties.

Figure 4. Illustration from our SR_VISION_LIMOSE data set

In our method, we opted for partitioning the data, where 50% was used for training, and 50% split between the validation and the testing. This allocation was intended to ensure optimal model training while maintaining a rigorous assessment of its performance on independent data. It achieved a balance between learning and assessment while ensuring that the model could be effectively generalized to new data. We used pre-trained ESRGAN models on natural images for several reasons (even though our main target was document images). Firstly, pre-trained models benefit from transfer learning, capturing a wide range of useful features applicable to various types of images. This technique significantly reduces the need for specific training data and computational resources. Natural images also possess textures/pixels, objects/backgrounds, etc., which helps the model learn very effectively how to generalize. This will let it support other kinds of images like document images out of the box with little extra adjustment. Natural images are easier to collect since extensive high-quality data sets of such can be found compared to specialized document images. Pre-trained models allow one to mitigate such challenges in smaller or worse quality data sets made for document images. The use of pre-trained models gave us an initial performance advantage because the model already understood many common image features. To make it even better, this model was fine-tuned by incorporating information from document images, which outperformed training everything from scratch. Deep-learning models are also computationally expensive to train from scratch, and it takes a lot of time. When one uses pre-trained models, the training time is significantly shortened, and far fewer resources are needed while still being able to realize excellent results. This way, the features learned by training on a large

base of images can be fine-tuned on our sets and adjusted to better suit document image specifics, combining broad generality with focused specialization. For training ESRGAN networks, parallel training was performed on five types of resulting images. The LL, LH, HL, and HH sub-bands, representing low frequencies, horizontal details, vertical details, and diagonal details, respectively, were bicubically interpolated to increase their resolution. Each type of image is then used to train a distinct ESRGAN network:

- ESRGAN_ORJ: trained on original image;
- ESRGAN_LL: trained on interpolated LL sub-band;
- ESRGAN_LH: trained on interpolated LH sub-band;
- ESRGAN_HL: trained on interpolated HL sub-band;
- ESRGAN_HH: trained on interpolated HH sub-band.

These networks are designed to produce high-resolution reproductions of input images, reducing reconstruction losses and optimizing the visual quality of generated images. The computational cost associated with driving multiple ESRGAN networks (one for each sub-band of wavelets as well as for the original image) can be high. This computational load results from the fact that each network requires significant GPU resources to be efficiently trained – especially with complex networks like ESRGAN (which are already very time-consuming). However, our method was eased by using pre-model training, which had significantly reduced the time and resources required for training. Using pre-existing models minimizes the need to recalculate parameters from zero. In addition, we used powerful GPUs, which allowed us to further reduce the processing time (thus, improving the overall efficiency of the process). We formed our models using the following configurations: the model was formed for a total of 20 epochs to ensure sufficient convergence and avoid overfitting. Batch size: to guarantee strong generalization and effective use of computing resources, a batch size of 32 was used for each training iteration. Learning rate: to ensure steady and progressive learning, we chose a starting learning rate of 0.001. This allowed us to balance the convergence speed and stability of the model while avoiding oscillations. This choice was based on proven practices for ESRGAN architecture. The loss function used was the mean square error (MSE); it calculates the difference in the mean square between the predicted values of the model and the actual values. This allows us to measure the gap between the model output and the reference image. MSE is commonly used in regression and super-resolution tasks, as it promotes rapid convergence while improving the accuracy of results.

The optimization function we used was Stochastic Gradient Descent (SGD). It is called "stochastic" because it randomly samples data at each iteration to estimate gradients, making training more effective on large data sets.
By adjusting these parameters, we obtained satisfactory super-resolution quality while maintaining an efficient and robust model formation.

Transfer learning is used to enhance the efficiency of training and the quality of the results by reusing the weights of a pre-trained model on a large data set.

Consequently, specialized ESRGAN networks improve the resolution and quality of images by leveraging the unique characteristics of each frequency sub-band.

3.3. Particle Swarm Optimization (PSO)

The PSO algorithm was inspired by the collective behavior of swarms of birds or schools of fish. It is an optimization algorithm that is implemented to solve certain types or formats of problems. A search space is explored using a population of elements (called 'particles' in PSO). Each particle corresponds to a possible solution of the problem and is defined by its location and speed in the search space. Each particle in the PSO algorithm updates its speed based on its individual experience (pbest solution) and collective memory value (gbest solution) within iterations. The particles are oriented to a new better position according to these two optimal values. The particles use their current speed and direction to move in the direction of the place that has produced the best results [29, 35, 43]. In the PSO algorithm, the velocity of particle "i" at time V_i^t depends on the velocity of the particle at time V_i^{t-1} ("pbest" and "gbest," respectively) representing the best position reached so far by an individual particle and by the whole swarm. Once these two best values are found, the particle updates its velocityV_i^t and position using the following equations:

$$V_i^t; = \underbrace{c1 \times V_i^{t-1}}_{\text{Current movement}} + \underbrace{c2 \times (pbest_i^t - X_i^t)}_{\text{Particle local search}} + \underbrace{c3 \times (gbest^t - X_i^t)}_{\text{Swarm influence}}$$

Here, X_i^t denotes the particle's actual location in the search area, where $c1$ is the inertia weight of the particle. When $c1$ is close to 0, the particle tends to follow its own best position more (*pbest*), which can lead to rapid convergence to a suboptimal (poor) solution. On the other hand, when $c1$ approaches 1, the search space is better explored, but the rate of convergence slows down. Constants $c2$ (individual confidence) and $c3$ (swarm confidence) influence the movement of the particle to a better position. c_2 represents the coefficient that determines the influence of the best individual position reached by a particle, while c_3 regulates the influence of the best overall position reached by the entire swarm. These values are usually selected randomly at the moment by a function like rand(0.1). In addition to *cmax*, a gravitational coefficient attracts particles to *gbest* (the best overall position) or *pbest* (the best personal position) to improve the search efficiency. However, an excessive value of *cmax* can lead to a divergence of the algorithm. It is crucial to emphasize that the optimal performance of the optimization is achieved when inertia and gravity work together. Therefore, the values of c1 *cmax* should not be selected independently for effective convergence. The following formula is then used to update particle position X_i^{t+1} at time $(t + 1)$:

$$X_i^{t+1}; = X_i^t \times v_i^t \tag{1}$$

A particle j may leave the originally set search space $[Xmin, Xmax]^D$ at any point during the swarm's development. According to this process, the value of the

nearest boundary point is then allocated. In practice, this means that Equation (1) must be changed to the following equation:

$$X_i^{t+1} = MIN(MAX(X_j^t \times V_j^t, X_{min}), X_{max}) \tag{2}$$

In addition, this mechanism is often supplemented by a gear change or by the replacement of the problematic component by its opposite (usually weighted by a coefficient less than 1 or by simply by canceling it). In our method, the parameters of the PSO algorithm were chosen to optimize the balance between the exploration and the exploitation of the solutions. Inertia factor $w = 0.5$ favors initial exploration, then rapid convergence. Coefficient $c_1 = 1$ ensures a balance between individual and collective particle searches, thus limiting their overdependence on past solutions. Coefficient $c_2 = 2$ encourages particles to focus more on the best overall solution, thus accelerating the convergence. A swarm of 15 particles and a range of initial velocities between -2 and 2 enable controlled exploration, while 50 iterations ensure rapid convergence and efficient optimization while minimizing computational costs. These choices offer an optimal compromise among exploration, exploitation, and computational efficiency. The PSO algorithm seeks to find optimal position $X^* = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5)$. Using these coefficients found by PSO, the results of the ESRGAN networks are merged to produce a final high-resolution image. This combination is expressed by the following formula:

$$\begin{aligned} I = {} & \alpha_1 \times I_{\text{ESRGAN\_ORJ}} + \alpha_2 \times I_{\text{ESRGAN\_LL}} + \alpha_3 \times I_{\text{ESRGAN\_LH}} \\ & + \alpha_4 \times I_{\text{ESRGAN\_HL}} + \alpha_5 \times I_{\text{ESRGAN\_HH}} \end{aligned} \tag{3}$$

where $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$, is a weighting generated by PSO algorithm, and I_{ESRGAN_ORG}, I_{ESRGAN_LL}, I_{ESRGAN_LH}, I_{ESRGAN_HL}, and I_{ESRGAN_HH} are the image results of the ESRGAN_ORG, ESRGAN_LL, ESRGAN_LH, ESRGAN_HL, and ESRGAN_HH models, respectively.

These coefficients play a crucial role in optimizing the balance between each sub-band's contribution and the original image, aiming to enhance the visual quality of the final image significantly. During the optimization process, different values of coefficients are examined to find those that maximize a chosen fidelity metric. The purpose of this was to maximize the Peak Signal to-Noise Ratio (PSNR) between the high-resolution output image and the reference image, indicating that the super-resolution process adequately keeps elaborately small detail and intactness in refining the quality images. The PSO in our PSO-WESRGAN method of super-resolution for document images is crucial. PSO is an optimization technique well-suited to our problem because it allows us to determine the optimal coefficients that weigh the different sub-bands of the super-resolved image. These coefficients are essential to balance the contribution of each sub-band and the original image to maximize the visual quality of the final image. The experiments demonstrated the effectiveness of our approach, showing significant improvements in terms of the sharpness, preservation of details, and overall visual quality of the super-resolved documentary images. These results

confirmed the ability of our method to meet the high demands of documentary image processing.

A PC running Windows 11 with an Intel i7-10750H CPU, 32 GB of RAM, 512 GB hard drive, and an Intel (R) UHDaphics card with an NVIDIA GeForce RTX 2060 graphics card was used to create our technique.

4. Evaluation and results

To assess the accuracy and effectiveness of our super-resolution technique for document images, it is essential to carefully examine its quality. We use PSNR and the Structural Similarity Index (SSIM) as two main indicators. PSNR is a popular measure to assess the quality of an over-resolved image against a high-resolution reference. It allows us to evaluate the performance of our super-resolution methods by quantifying the noise level in the super-resolution image compared to the reference. Calculated using the logarithm of the maximum pixel range (typically 255 for RGB images) divided by the square root of the mean deviation between the corresponding pixels in the super-resolution and reference images, PSNR is expressed in decibels (dB). A higher PSNR value indicates better super-resolution image quality.

$$PSNR(I_1, I_2) = 10 \log \left(\frac{255^2}{\sqrt{MSE(I_1, I_2)}} \right) \tag{4}$$

where:

$$MSE(I_1, I_2) = \left(\frac{1}{RC} \sum_{X=0}^{R-1} \sum_{Y=0}^{C-1} ||I_1(x, y) - I_2(x - y)||^2 \right. \tag{5}$$

The SSIM measure examines image quality in three key aspects: luminance similarity (L), contrast similarity (C), and structure similarity (S). These factors are combined to provide an overall assessment of the structural similarity between two images (I_1 and I_2). SSIM is considered more advanced than PSNR because it takes into account how humans perceive images. It is often considered a better indicator of perceived image quality and is calculated by the following formula:

$$L(I_1, I_2) = \frac{2u_{I1}uI2 + c1}{u_{I1}^2 + uI2^2 + c1} \tag{6}$$

$$C(I_1, I_2) = \frac{2\sigma_{I1}\sigma_{I2} + c1}{\sigma_{I1}^2 + \sigma_{I2}^2 + c1} \tag{7}$$

$$S(I_1, I_2) = \frac{2\sigma_{I1I2} + c2/2}{\sigma_{I1}\sigma_{I2} + c2/2} \tag{8}$$

$$SSIM(I_1, I_2) = L(I_1, I_2)^\alpha \times C(I_1, I_2)^\beta \times S(I_1, I_2)^\gamma \tag{9}$$

By setting weights *alpha*, *beta*, and *gamma*, we simplify the formula. The $c1$ and $c2$ parameters stabilize the division by keeping the denominator high. The average and standard deviations of the image are represented by u and *sigma*, respectively.

To show the importance of each wavelet sub-band in the super-resolution process, we performed several comparative studies using various sub-band configurations. We first simply used the LL sub-band. This allowed us to watch how well the model performed when a high-resolution image was predicted using only the global information from the low-resolution image. Then, we included the HH sub-band, which contains diagonal details. However, this last step did not provide a significant improvement over the use of the LH and HL sub-bands, as the details of the other bands can sometimes overlap. This redundancy may explain why the PSNR improvement was not as significant as when we focused on the LH sub-band alone. Finally, we used all of the available sub-bands, including LL, LH, HL, and HH; this allowed us to integrate all of the information contained in these different sub-bands for super-resolution. Each of these sub-bands contributes to capturing specific image details, whether coarse or fine and horizontal or vertical. These steps show how important it is to choose the right sub-bands or the right weighting of sub-bands according to the information they contain to get the best super-resolution results. This demonstrates the strategic and methodical approach needed to optimize the super-resolution process using PSO. Table 1 shows the average PSNR values for each LIMOSE-based configuration.

Table 1

Representative table of PSNR values on LIMOSE database

Configuration	PSNR [dB]
ORJ + LL	32,84903
ORJ + LL + LH	33,61347
ORG+LL + LH + HL	33,63208
ORG + LH	33,64842
ORG + LH + HL	33,50335
ORG + LH + HL+ HH	32,9637
Our method (ORG + LL + LH + HL + HH)	33,65108

We observe a significant improvement in PSNR as more wavelet sub-bands are included in the super-resolution model. The configuration that uses all of the sub-bands shows the best performance in terms of predicted image quality on this specific database. The curves in Figures 5 and 6 show the evaluation of our approach with the different methods using the PSNR and SSIM measures. It should be noted that the specific PSNR values for each method depend on the parameters used, the data set, and the specific images to be evaluated. The PSNR values measure the quality of the super-resolved image compared to the high-resolution reference image; higher values indicate better image quality in terms of pixel fidelity.

From Figure 5 and Table 2, we can see that our method has the highest PSNR score, indicating better super-resolved image quality in terms of fidelity to the reference image details. ESRGAN comes in second, closely followed by SRGAN, SRCNN and Bell, Bilinaire, Bspline, Contour stencils, Gaussien, Hanning, Hermite, Lanczos, Neighbor, and Mitchell obtaining the lowest scores. It is important to note that PSNR is a quantitative metric that does not take into account certain aspects of the visual

quality perceived by the human eye. It is therefore recommended to complete the evaluation with SSIM and human observation to better evaluate the image quality. Figure 6 shows that our approach tends to obtain better SSIM values, indicating better structural similarity between super-resolved images and high-resolution reference images compared to the other methods.

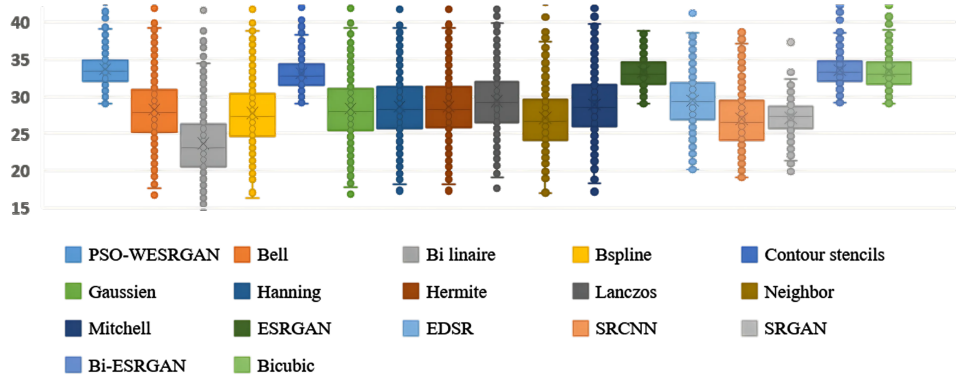

Figure 5. Representative PSNR curves obtained by applying different methods on LIMOSE database

Table 2
Representative table of PSNR and SSIM values on LIMOSE database

Methods	PSNR (dB)	SSIM
Our method	33,65108	0,849795998
Bell	29,42585	0,797391499
Bi linaire	25,81925	0,796938341
Bspline	29,04785	0,698770523
Contour stencils	27,569125	0,788846099
Gaussien	29,523	0,753854316
Hanning	29,661825	0,798700025
Lanczos	29,678475	0,795954727
Neighbor	28,47185	0,796712643
Mitchell	29,8833	0,798971884
Bicubic	33,3849918	0,799521700
ESRGAN	33,30121989	0,761439613
EDSR	29,43117478	0,802635148
SRCNN	26,94576659	0,805519466
SRGAN	27,09932406	0,762026427
Bi-ESRGAN	33,50647544	0,840099273

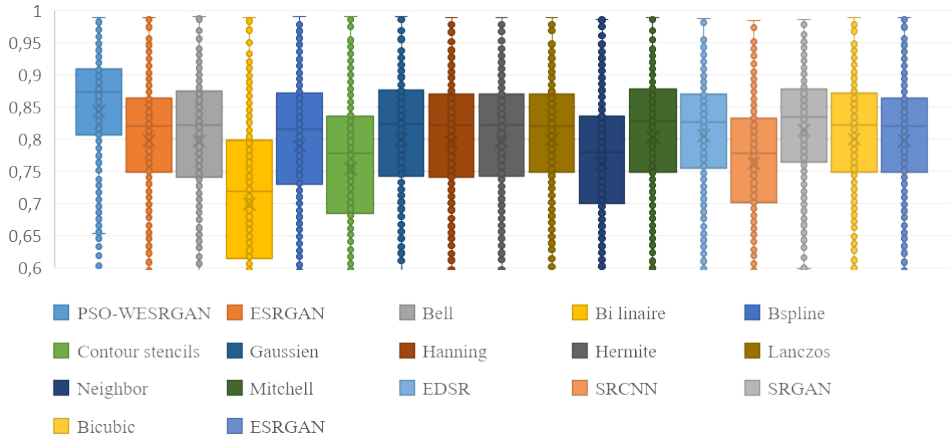

Figure 6. Representative SSIM curves obtained by applying different methods on LIMOSE database

Figures 7 and 8 present an example of how our method performs when compared to classic techniques (Bicubic, Biliniar, Lanactoz, and Nearest Neighbor) and deep-learning methods (SRCNN, EDSR, SRGAN, ESRGAN).

Deep-neural-network-based image super-resolution techniques ESRGAN, SRGAN, SRCNN, and EDSR vary in their architectures and image quality performance. Figures 5, 6, 7, and 8 illustrate how our method more accurately traced the characters and defined their shapes.

Figures 5, 6, and 7 show that other methods demonstrate good performance in image accuracy and quality but have limitations in rendering extremely fine details or complex textures – sometimes resulting in the loss of subtleties. Comparing the results of our algorithm with classical methods such as Bicubic, Bilinear, Lanczos, and Nearest Neighbor showed that PSO-WESRGAN outperforms these methods in terms of PSNR and SSIM, thus indicating better image quality and preservation of image structures. Visually, our method produces images that are sharper and more faithful to the details of the reference image, with fewer artifacts and better readability of documents. SRCNN and EDSR focus on learning nonlinear relationships, showing good performance but with limitations for very fine details. ESRGAN and SRGAN significantly improve super-resolved image quality but may be sensitive to artifacts and irregularities in generated images. Due to the complexity of generative adversarial networks (GANs), there may be cases where ESRGAN and SRGAN results contain visible defects, such as blurred contours, halos around objects, and artificial textures. These artifacts can affect the perceived quality of an image and make the results less natural.

To evaluate our method effectively, we measured its impact on the accuracy of OCR systems by comparing it with other super-resolution methods. This comparison was performed using a set of low-resolution documentary images processed by our

approach as well as competing techniques (SRGAN, SRCNN, and other reference methods). The super-resolved images were then analyzed by a standard OCR tool.

Figure 7. Visual comparison between our method and traditional methods: (1) LR image; (2) PSO_WESRGAN; (3) Bicubic; (4) Biliniar; (5) Lanactoz; (6) Nearest Neighbor

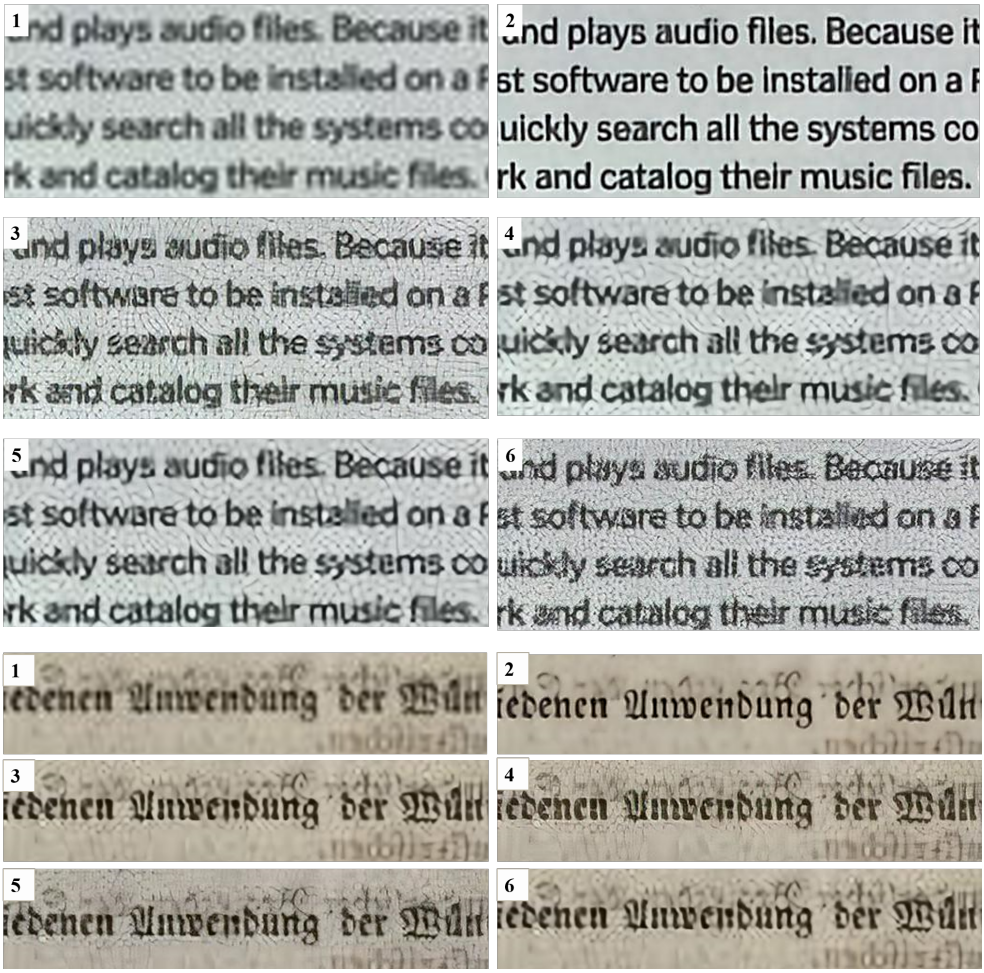

Figure 8. Comparison of efficacy of our method compared to deep-learning methods: (1) LR image; (2) PSO_WESRGAN; (3) ESRGAN; (4) EDSR; (5) SRGAN; (6) SRCNN

Table 3
Comparison using OCR metrics

Methods	OCR%
Our method	98.89
ESRGAN	79.30
EDSR	72.73
SRGAN	78.55
SRCNN	58.83

The results obtained (illustrated in Figure 9 and Table 3) show that our method significantly improves image quality, resulting in a significant reduction of OCR errors. This improvement demonstrates the effectiveness of our method to preserve document details, making automatic text extraction more reliable.

Figure 9. Comparison using OCR metrics: (1) HR image; (2) PSO_WESRGAN; (3) ESRGAN; (4) EDSR; (5) SRGAN; (6) SRCNN

To test the performance of our method on other types of images, we performed a full visual comparison (Figures 10, 11, 12, and 13) and a quantitative evaluation (Table 4) of our method compared to the best-known approaches in the literature (which were chosen for their popularity and performance in the field of image super-resolution). We chose well-known and widely used databases in the field of super-resolution, such as Set5, Set14, BSD100, and Urban. These data sets contain a wide variety of natural images covering different types of scenes and resolutions and are used to test the robustness and effectiveness of super-resolution methods on more-general images. The diversity of these sets allows us to evaluate the performance of our method in various contexts and compare its results with those of other approaches established in the field.

Our super-resolution technique excelled in our extensive comparative study, outperforming well-known models on various key data sets. Our approach was distinguished by excellent detail preservation, strong structural similarity to other models, and superior PSNR and SSIM values on the Set5 and Set14 data sets. Similarly, our

method outperformed BSD100, delivering outstanding results in clarity and visual integrity. However, it was on the Urban100 data set that our solution shone, outperforming all other comparative models and demonstrating its exceptional ability to process complex images in urban environments. These promising results underline the resilience and flexibility of our model in a variety of situations, positioning our method as a first-choice option for image super-resolution in diverse contexts.

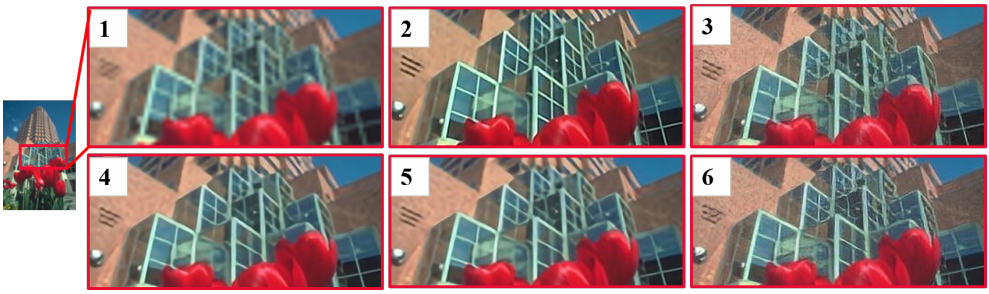

Figure 10. Visualization results for 4× super-resolution on image img_095 from BSD100 data set

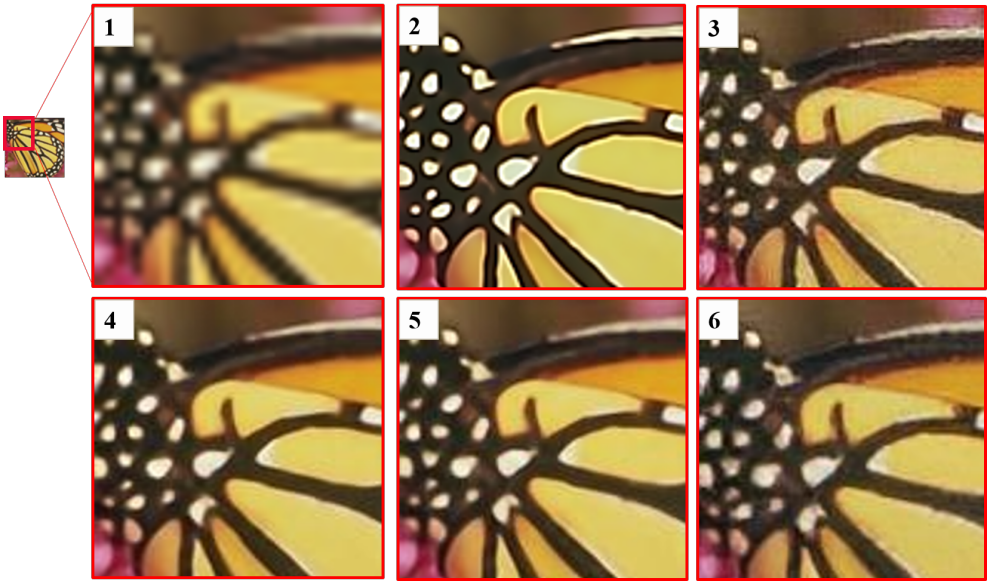

Figure 11. Visualization results for 4× super-resolution on image img_003 Set5 data set

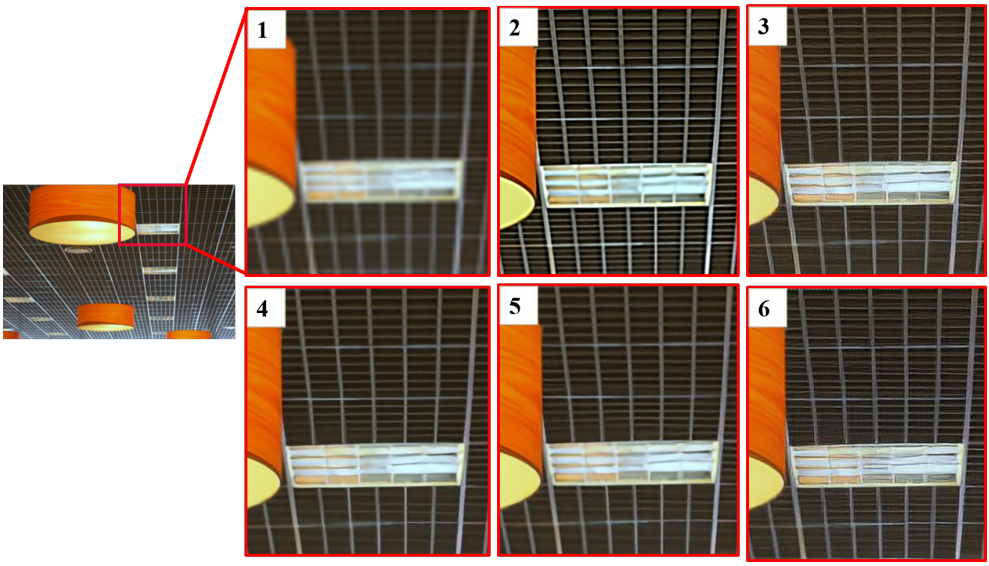

Figure 12. Visualization results for 4× super-resolution on image img_044 Urban data set

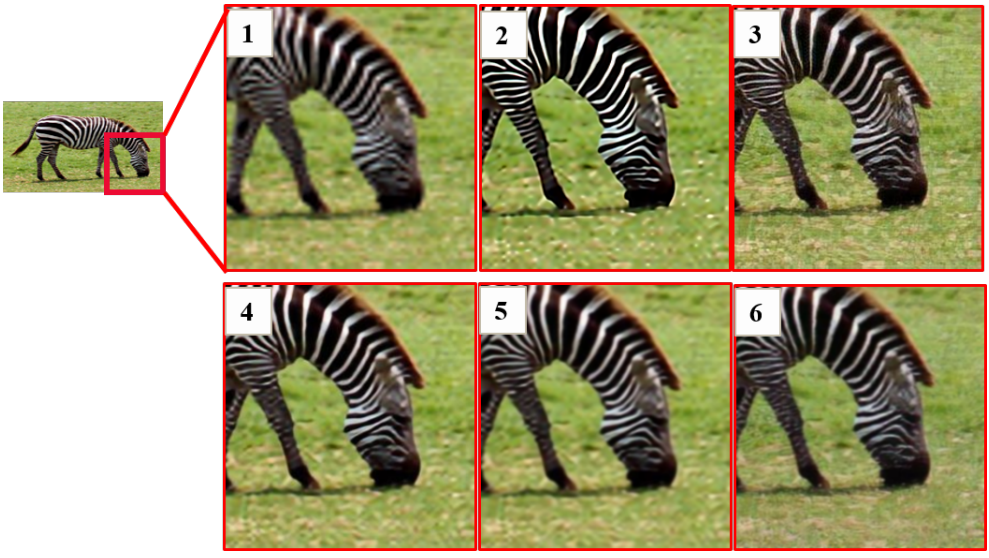

Figure 13. Visualization results for 4× super-resolution on image img_014 Set14 data set

Table 4
PSNR and SSIM values using different models on benchmark (Set5/Set14/BSD100/Urban100) data sets for 4× super resolution

Method	Data Set	Set5	Set14	BSD100	Urban
SRCNN	291	30.48 0.86	27.50 0.75	26.90 0.71	24.52 0.72
EDSR	DIV2K	32.46 0.89	28.80 0.78	27.71 0.74	26.64 0.80
RCAN	DIV2K	32.63 0.90	28.87 0.78	27.77 0.74	26.82 0.80
ESRGAN	DF2K	32.73 0.90	28.99 0.79	27.85 0.74	27.03 0.81
MDRN	DIV2K	32.33 0.89	28.75 0.78	26.43 0.79	27.66 0.73
EdgeSRGA	DIV2K	31.72 0.88	28.30+ 0.77	27.35 0.72	25.4 0.76
HAT	DIV2K	32.92 0.90	29.15 0.79	27.97 0.75	27.87 0.83
DRCT-L [17]	DF2K	33.37 0.90	29.54 0.80	28.16 0.75	28.70 0.85
HMA [4]	DF2K ImageNet	33.38 0.90	29.51 0.80	28.13 0.75	28.69 0.85
CPAT+ [40]	DF2K ImageNet	33.19 0.9069	29.51 0.7991	28.04 0.75	28.22 0.8408
CFAT [31]	DIV2K Flickr2K	33.19 0.90	29.34 0.79	29.30 0.79	28.11 0.83
SAFMN [37]	DIV2K Flickr2Kt	32.18 0.89	28.60 0.78	27.58 0.73	25.97 0.78
SVAN [9]	DIV2K Flickr2Kt	31.76 0.88	28.30 0.7736	27.41 0.72	25.56 0.76
Extracter-rec [32]	CUFED5	27.29 0.811	30.02 0.816	26.04 785	28.09 0.782
DSRNet [39]	DIV2K	31.71 0.88	28.38 0.7760	27.43 0.73	25.65 0.76
SwinIR	DIV2K +Flickr2K	32.93 0.9043	29.15 0.7958	27.95 0.7494	27.56 0.8273
PSO- WESRGAN	DIV2K+ SR _LIMOSE	33,58 0.92	30,21 0.798	29,55 0.81	28,25 0.88

5. Conclusion

In conclusion, our article proposes an innovative method for the super-resolution of documentary images, integrating wavelet transformation, deep-learning transfer, and optimization by the PSO algorithm. This approach has allowed us to preserve the

subtle details and layout of documents while optimizing parameters to significantly improve the visual quality of images. The experimental results show significant gains in terms of fidelity, detail, and readability.

However, several limitations should be noted. First, the approach is computationally demanding. ESRGAN networks require high GPU power for the training and inference steps, and the PSO algorithm (by exploring a large number of parameters) further intensifies these requirements. This constraint can make it difficult to apply our method on devices with low computing capacities (such as embedded or portable systems). Also, images with very low resolution, high noise, or large artifacts may not be improved sufficiently. ESRGAN networks and PSO optimization may reconstruct details incorrectly.

These limitations call for a thorough reflection to adapt and improve our method. Future prospects include the development of lighter architectures to reduce computational costs. Integrating more-efficient PSO variants to limit the complexity of the optimization and developing robust data preprocessing and enhancement techniques would improve the performance on low-quality images. Furthermore, a thorough validation on more-varied data sets would allow the application of the method to be extended.

Acknowledgements

No particular grant was given for this research by governmental, private, or nonprofit funding organizations.

References

[1] Akima H.: A method of bivariate interpolation and smooth surface fitting for irregularly distributed data points, *ACM Transactions on Mathematical Software (TOMS)*, vol. 4(2), pp. 148–159, 1978. doi: 10.1145/355780.355786.

[2] Behjati P., Rodriguez P., Fernández C., Hupont I., Mehri A., Gonzàlez J.: Single image super-resolution based on directional variance attention network, *Pattern Recognition*, vol. 133, 108997, 2023. doi: 10.1016/j.patcog.2022.108997.

[3] Chang Y., Chen G., Chen J.: Pixel-Wise Attention Residual Network for Super-Resolution of Optical Remote Sensing Images, *Remote Sensing*, vol. 15(12), 3139, 2023. doi: 10.3390/rs15123139.

[4] Chen X., Wang X., Zhou J., Qiao Y., Dong C.: Activating more pixels in image super-resolution transformer. In: *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 22367–22377, 2023. doi: 10.1109/cvpr52729.2023.02142.

[5] Deng L., Zhang Y., Yang X., Huang S., Wang J.: Meta-Learning Multi-Scale Radiology Medical Image Super-Resolution, *Computers, Materials & Continua*, vol. 75(2), pp. 2671–2684, 2023. doi: 10.32604/cmc.2023.036642.

[6] Dohyun K., Joongheon K., Junseok K., Tae-Hyung K.: Depth-controllable very deep super-resolution network. In: *International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2019. doi: 10.1109/ijcnn.2019.8851874.

[7] Dong C., Loy C.C., He K., Tang X.: Learning a Deep Convolutional Network for Image Super-Resolution. In: D. Fleet, T. Pajdla, B. Schiele, T. Tuytelaars (eds.), *Computer Vision – ECCV 2014. 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part IV*, pp. 184–199, 2014. doi: 10.1007/978-3-319-10593-2_13.

[8] El-Shafai W., Mohamed E.M., Zeghid M., Ali A.M., Aly M.H.: Hybrid Single Image Super-Resolution Algorithm for Medical Images, *Computers, Materials and Continua*, vol. 72(3), pp. 4879–4896, 2022. doi: 10.32604/cmc.2022.028364.

[9] Fan Q., Wu C., Hu S., Wu X., Wang X., Hu J.: Efficient Image Super-Resolution via Symmetric Visual Attention Network. In: *2024 International Joint Conference on Neural Networks (IJCNN)*, 2024. doi: 10.1109/ijcnn60899.2024.10650362.

[10] Gallivan K., Grimme G., Van Dooren P.: Arational Lanczos algorithm for model reduction, *Numerical Algorithms*, vol. 12, pp. 33–63, 1996. doi: 10.1007/bf02141740.

[11] Getreuer P.: Contour stencils for edge- adaptive image interpolation. In: M. Rabbani, R.L. Stevenson (eds.), *Visual Communications and Image Processing 2009*, vol. 7257, 2009. doi: 10.1117/12.806014.

[12] Getreuer P.: Image interpolation with contour stencils, *Image Processing on Line*, vol. 1, pp. 70–82, 2011. doi: 10.5201/ipol.2011.g_iics.

[13] Gordon W.J., Riesenfeld R.F.: B-spline curves and surfaces. In: R.E. Barnhill, R.F. Riesenfeld (eds.), *Computer Aided Geometric Design*, pp. 95–126, Academic Press, 1974. doi: 10.1016/b978-0-12-079050-0.50011-4.

[14] Grandke T.: Interpolation algorithms for discrete Fourier transforms of weighted signals, *IEEE Transactions on Instrumentation and Measurement*, vol. 32(2), pp. 350–355, 1983. doi: 10.1109/tim.1983.4315077.

[15] Hall C.A.: Bicubic interpolation over triangles, *Indiana University Mathematics Journal*, vol. 19(1), 1970. doi: 10.1512/iumj.1970.19.19001.

[16] He Z., Chen D., Cao Y., Yang J., Cao Y., Li X., Tang S., *et al.*: Single image super-resolution based on progressive fusion of orientation-aware features, *Pattern Recognition*, vol. 133, 109038, 2023. doi: 10.1016/j.patcog.2022.109038.

[17] Hsu C.C., Lee C.M., Chou Y.S.: DRCT: Saving Image Super-Resolution away from Information Bottleneck. In: *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 6133–6142, 2024. doi: 10.1109/cvprw63382.2024.00618.

[18] Kezzoula Z., Gaceb D., Akli Z., Kahouli A., Titoun A., Touazi F.: Bi-ESRGAN: A New Approach of Document Image Super-Resolution Based on Dual Deep Transfer Learning. In: M. Salem, J.J. Merelo, P. Siarry, R. Bachir Bouiadjra, M. Debakla, F. Debbat (eds.), *Artificial Intelligence: Theories and Applications. First International Conference, ICAITA 2022, Mascara, Algeria, November 7–8, 2022, Revised Selected Papers*, pp. 110–122, Springer, 2022. doi: 10.1007/978-3-031-28540-0_9.

[19] Lai W.S., Huang J.B., Ahuja N., Yang M.H.: Fast and accurate image super-resolution with deep laplacian pyramid networks, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41(11), pp. 2599–2613, 2018. doi: 10.1109/tpami.2018.2865304.

[20] Ledig C., Theis L., Huszár F., Caballero J., Cunningham A., Acosta A., Aitken A., *et al.*: Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 105–114, IEEE, 2017. doi: 10.1109/cvpr.2017.19.

[21] Liang J., Cao J., Sun G., Zhang K., Van Gool L., Timofte R.: SwinIR: Image restoration using swin transformer. In: *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, pp. 1833–1844, 2021. doi: 10.1109/iccvw54120.2021.00210.

[22] Lim B., Son S., Kim H., Nah S., Lee K.M.: Enhanced Deep Residual Networks for Single Image Super-Resolution. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1132–1140, 2017. doi: 10.1109/cvprw.2017.151.

[23] Luo J., Han L., Gao X., Liu X., Wang W.: SR-FEINR: Continuous Remote Sensing Image Super-Resolution Using Feature-Enhanced Implicit Neural Representation, *Sensors*, vol. 23(7), 3573, 2023. doi: 10.3390/s23073573.

[24] Mahapatra D., Bozorgtabar B., Garnavi R.: Image super-resolution using progressive generative adversarial networks for medical image analysis, *Computerized Medical Imaging and Graphics*, vol. 71, pp. 30–39, 2019. doi: 10.1016/j.compmedimag.2018.10.005.

[25] Mao Y., Zhang N., Wang Q., Bai B., Bai W., Fang H., Liu P., *et al.*: Multi-level Dispersion Residual Network for Efficient Image Super-Resolution. In: *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1660–1669, 2023. doi: 10.1109/CVPRW59228.2023.00167.

[26] Mitchell D.P., Netravali A.N.: Reconstruction filters in computer-graphics, *ACM SIGGRAPH Computer Graphics*, vol. 22(4), pp. 221–228, 1988. doi: 10.1145/378456.378514.

[27] Narasimhan S.V., Basumallick N., Veena S.: *Introduction to Wavelet Transform: A Signal Processing Approach*, Alpha Science, 2011.

[28] Okabe M.: Explicit interpolation formulas for the Bell triangle, *Computer Methods in Applied Mechanics and Engineering*, vol. 117(3–4), pp. 411–421, 1994. doi: 10.1016/0045-7825(94)90126-0.

[29] Olsson A.E. (ed.): *Particle Swarm Optimization: Theory Techniques and Applications*, Nova, 2011.

[30] Othman G., Zeebaree D.Q.: The applications of discrete wavelet transform in image processing: A review, *Journal of Soft Computing and Data Mining*, vol. 1(2), pp. 31–43, 2020.

[31] Ray A., Kumar G., Kolekar M.H.: CFAT: Unleashing Triangular Windows for Image Super-resolution. In: *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 26120–26129, 2024. doi: 10.1109/cvpr52733.2024.02468.

[32] Reyes-Saldana E., Rivera M.: EXTRACTER: Efficient Texture Matching with Attention and Gradient Enhancing for Large Scale Image Super Resolution, 2023. doi: 10.2139/ssrn.5347487.

[33] Salvetti F., Mazzia V., Khaliq A., Chiaberge M.: Multi-Image Super Resolution of Remotely Sensed Images Using Residual Attention Deep Neural Networks, *Remote Sensing*, vol. 12(14), 2020. doi: 10.3390/rs12142207.

[34] Schumer M.: Interpolation of a Gaussian-Markov process (Corresp.), *IEEE Transactions on Information Theory*, vol. 16(1), pp. 75–77, 1970. doi: 10.1109/tit.1970.1054397.

[35] Shi Y.: Particle swarm optimization, *IEEE Connections*, vol. 2(1), pp. 8–13, 2004.

[36] Smith P.R.: Bilinear interpolation of digital images, *Ultramicroscopy*, vol. 6(2), pp. 201–204, 1981. doi: 10.1016/0304-3991(81)90061-9.

[37] Sun L., Dong J., Tang J., Pan J.: Spatially-adaptive feature modulation for efficient image super-resolution. In: *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 13190–13199, 2023. doi: 10.1109/iccv51070.2023.01213.

[38] Tao Y., Muller J.P.: Super-Resolution Restoration of Spaceborne Ultra-High-Resolution Images Using the UCL OpTiGAN System, *Remote Sensing*, vol. 13(12), 2021. doi: 10.3390/rs13122269.

[39] Tian C., Zhang X., Zhang Q., Yang M., Ju Z.: Image super-resolution via dynamic network, *CAAI Transactions on Intelligence Technology*, vol. 9(4), pp. 837–849, 2024. doi: 10.1049/cit2.12297.

[40] Tran D.P., Hung D.D., Kim D.: Channel-Partitioned Windowed Attention And Frequency Learning for Single Image Super-Resolution, *arXiv preprint arXiv:240716232*, 2024. doi: 10.48550/arXiv.2407.16232.

[41] Tsai R.Y., Huang T.S.: Multiframe image restoration and registration. In: T.S. Huang (ed.), *Advances in Computer Vision and Image Processing*, vol. 1, pp. 317–339, JAI Press, Greenwich CT, 1984.

[42] Urbina Ortega C., Quevedo Gutiérrez E., Quintana L., Ortega S., Fabelo H., Santos Falcón L., Marrero Callico G.: Towards Real-Time Hyperspectral Multi-Image Super-Resolution Reconstruction Applied to Histological Samples, *Sensors*, vol. 23, 4, 2023. doi: 10.3390/s23041863.

[43] Walker B. (ed.): *Particle Swarm Optimization (PSO). Advances in Research and Applications*, Nova, 2017.

[44] Wang X., Yu K., Wu S., Gu J., Liu Y., Dong C., Qiao Y., *et al.*: ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks. In: L. Leal-Taixé, S. Roth (eds.), *Computer Vision – ECCV 2018 Workshops. Munich, Germany, September 8–14, 2018, Proceedings, Part V*, pp. 63–79, Springer, Cham, 2019. doi: 10.1007/978-3-030-11021-5_5.

[45] Wang Z., Cun X., Bao J., Zhou W., Liu J., Li H.: Uformer: A General U-Shaped Transformer for Image Restoration. In: *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 17662–17672, 2022. doi: 10.1109/CVPR52688.2022.01716.

[46] Zhang K., Zuo W., Zhang L.: Deep Plug-And-Play Super-Resolution for Arbitrary Blur Kernels. In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1671–1681, 2019. doi: 10.1109/cvpr.2019.00177.

[47] Zhang Y., Li K., Li K., Wang L., Zhong B., Fu Y.: Image Super-Resolution Using Very Deep Residual Channel Attention Networks. In: V. Ferrari, M. Hebert, C. Sminchisescu, Y. Weiss (eds.), *Computer Vision – ECCV 2018. 15th European Conference, Munich, Germany, September 8–14, 2018, Proceedings, Part VII*, pp. 286–301, Springer, Cham, 2018. doi: 10.1007/978-3-030-01234-2_18.

[48] Zhang Y., Tian Y., Kong Y., Zhong B., Fu Y.: Residual Dense Network for Image Super-Resolution. In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2472–2481, IEEE, 2018. doi: 10.1109/cvpr.2018.00262.

Affiliations

Zakia Kezzoula
University M'Hamed Bougara of Boumerdes, LIMOSE Laboratory, Algeria,
z.kezzoula@univ-boumerdes.dz

Djamel Gaceb
University M'Hamed Bougara of Boumerdes, LIMOSE Laboratory, Algeria,
d.gaceb@univ-boumerdes.dz

Received: 30.07.2024
Revised: 23.01.2025
Accepted: 4.05.2025