Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

Single Image Super-Resolution Using Deep Ensemble Learning

Faraz Mohammadian jadval ghadam[1], Sattar Hashemi[2,*], Karamollah Bagherifard [3], Seyed Mehdi Hazrati Fard[4], Samad Nejatian[5]

[1] Ph.D. candidate, Department of Computer Engineering, Yasuj Branch, Islamic Azad University, Yasuj, Iran. Email: faraz.mohammadian@yums.ac.ir
[2*] Ph.D. Professor of Artificial Intelligence, Shiraz University,Shiraz,Fars, Iran. Email: s_hashemi@shirazu.ac.ir
[3] Ph.D. Assistant professor, Department of Computer Engineering, Yasuj Branch, Islamic Azad University, Yasuj, Iran. Email: ka.bagherifard@iau.ac.ir
[4] Ph.D. Faculty of Graduate Studies, Department of Computer Engineering,University of Victoria,Canada. Email: smhazrati@uvica.ca
[5] Ph.D. Assistant professor, Department of Computer Engineering, Yasuj Branch, Islamic Azad University, Yasuj, Iran. Email: Sa.Nejatian@iau.ac.ir
***Corresponding Author:** Sattar Hashemi

***Abstract*—**The advancement in the digital world has made it possible to take high-quality images with even cellphones, but the size of the captured photos is a challenging issue since transferring and storing them are costly. Compression algorithms are widely used in industry and aerial imagery since some details are lost while compressing the images. These lost details are needed in some cases and applications. To address this issue, this paper proposes an ensemble deep learning method that shares information between different deep learning models to recover the original image. These models are employed to recover high-quality versions of images, known as super-resolution. We benefited from the Single Image technique and used one image as the input of the deep networks. To perform the ensemble learning, a CNN was used to integrate the outputs of different GAN models (LapSRN, SRResNet, RESNeXt, SRCNN, and ESCPN) to generate the high-resolution image. The proposed method could achieve intriguing results to increase the recovery quality compared to the rival methods in this realm.

Keywords*: Super_resolution - Single Image Super_resolution – GAN - Ensemble Deep Learning*

I. INTRODUCTION

There has always been a demand for high-quality images in both commercial and scientific areas. However, taking high-quality images is either impossible or expensive in some cases. For instance, mobile phones, mainly made for calls, may also have cameras to let users take

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

photos; however, producing phones with big cameras is not favorable. In addition, in scientific applications such as satellite image processing and security, the transmission of high-quality images requires a high bandwidth network, which is very expensive and sometimes impossible when there is a large amount of data for transition. As another example, in medical cases, the taken image is not as high-quality as the medical staff expects it. These demands, among many others, have motivated machine vision and artificial intelligence (AI) experts to turn low-quality images into high-quality ones.One of the attempts to solve this problem is super-resolution (SR). SR is a software method in which one or more low-resolution (LR) images are taken as input to make a high-resolution (HR) image [1]. This method has been employed in different areas such as medical imaging, satellite image processing, video enhancement, security, and computer vision. In other words, super-resolution can improve both perceptual image quality and computer vision [2]. The examination of the literature shows that some classical SR methods such as edge-based [3], patch-based [4], sparse representation [5], prediction-based [6], and statistical methods [7] have been used extensively for SR. Recent models in SR leverage the ability of the vector image system to improve the quality of converted images. The main objective of the SR technique is to increase the spatial resolution of images. In order to construct a higher resolution image, Single Image Super-resolution (SISR) requires an image as the input; however, SR needs multiple images to create a higher resolution image that is not affordable in most cases [2]. As a result, SISR can address the problem of requiring several images and significantly facilitate the process when more than one photo is not easily accessible.

Today, the field of deep learning is used in all aspects of human life [8]. .Deep learning has been widely used to optimize the SISR method in recent years. Another advancement in classification technology is ensemble learning, combining several individual models to obtain better generalization performance, and has had many applications[9,10]. Currently, deep learning models with multilayer processing architecture are showing better performance than the shallow or traditional classification models [1]. Ensemble Learning utilizes an array of learning algorithms to achieve efficient classification results. First, a set of classifiers will be trained with an identical dataset during the training sessions. The results are then combined to reach an optimized classification. The employment of one classifier can result in bias, while the employment of different classifiers through ensemble learning can minimize the bias. There are different ensemble learning types. Bootstrap aggregating is used to enhance the classification result and helps to minimize the overfitting [7]. Bootstrap aggregation (also called bagging) combines prediction outputs of models educated independently. One of the examples of this category is the forest algorithm [6]. The second type of ensemble learning is boosting, which combines a group of classifiers to shape a more robust classifier. A bucket of models is also used to select the most efficient model [7]. The proposed method aims to

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

improve the quality of an image's resolution. To obtain improved results, the researchers have used ensemble learning to minimize the weaknesses of each single model by using different data sources. The present paper suggests integrating the SISR model with deep ensemble learning to optimize the process of single image super-resolution. The deep ensemble learning model combines the advantages of both deep learning models as well as ensemble learning such that the final model has better generalization performance. This study shows how the combination of these two technologies can result in achieving high-quality images.The rest of this paper is organized as follows. Section 2 provides information on related studies in the literature. The proposed method is discussed in Section 3, which provides details on the innovative methods. Section 4 includes the experimental results. Section 5 presents the achieved results and compares them with state of the art. Finally, Section 6 concludes the paper and discusses the possibilities for further studies.

II. RELATED WORKS

Creating HR images from LR images has always been one of the concerns of researchers [15-18]. The SISR problem is inherently an underdetermined problem [13] because an LR image can be converted into many HR images in better quality, and the problem contains several solutions [9]. Many SISR methods have been developed in computer vision [11, 12]. Among them, interpolation methods are easy to implement and widely adopted [15]. However, these linear models have limited representation power and often generate blurry HR outputs. Sparsity-based techniques [5, 14] have been recently developed to enhance linear models with rich image priors.The Bicubic SR method [6] provides a testbed that helps you analyze different models simultaneously. Bicubic SR attempts seek to fit a surface among corner pixels. In this method, the vertical, horizontal, and diagonal derivate, as well as intensity values, are computed. This method can help us identify the gap between the unsupervised and supervised frameworks. It has been employed extensively since it imposes low complexity and relatively acceptable results.

Since the emergence of GAN (generative adversarial network) [19], several methods have been proposed to improve the results. In SISR, SRGAN (super-resolution generative adversarial network) has proved to be an efficient method [20-22], but its ability to improve the quality of images is not acceptable in some cases. On the other hand, LapGAN (Laplacian generative adversarial network) [21, 23] is an efficient network for converting illusions into real images. A network consists of several layers used to increase the number of images. Therefore, it has been able to process more images and provide better results. By merging SRGAN and LapGAN, a new model called LapSRN (Laplacian pyramid super-resolution network) was created [24]. LapSRN has a high capability in HR production and acceptable

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

quality. This method demonstrates better results than the other two methods. Figure 1 depicts the overall structure of LapSRN.

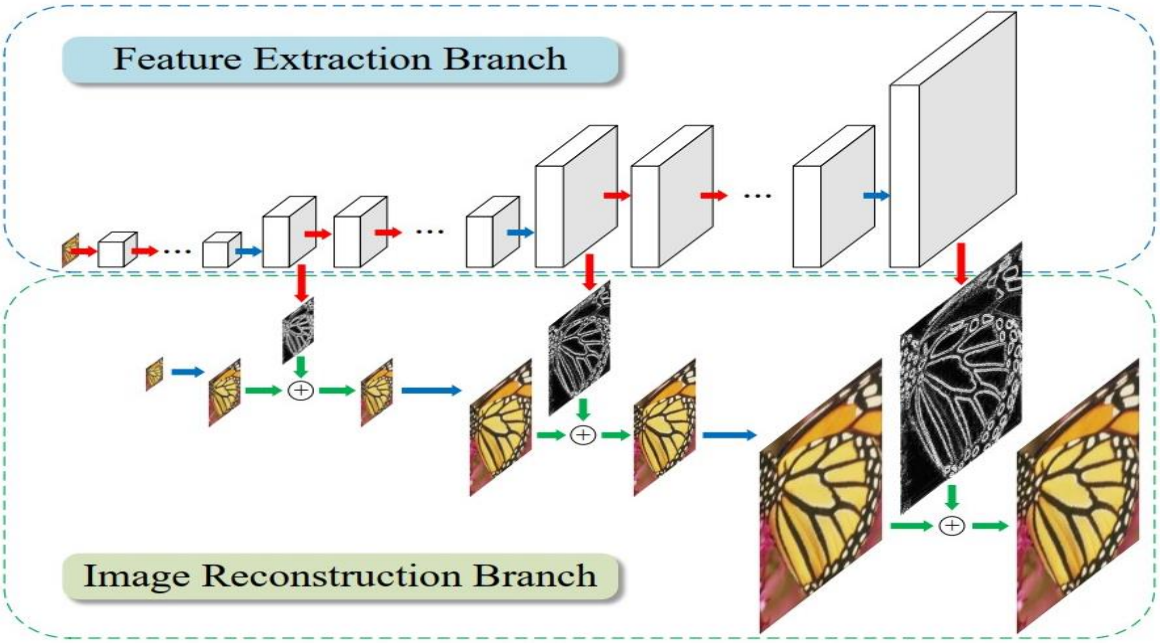

Figure 1. LapSRN [24].

HR images are expensive and have a lower signal-to-noise ratio in areas such as aerial or medical imagery. Various models using machine learning methods have been offered so far. Deep networks have achieved significant results in this field with two-dimensional and three-dimensional data. On the other hand, deep networks require the reflection of the nature of data under consideration in the network structure or the objective function in the form of regularization. The use of different deep networks such as Inception, ResNet (residual neural network) [16], and DCCN (dense connected cascade network) [13] in combination with target functions based on perceptual terms to achieve higher quality images requires study and research [25, 26].

The SRResNet, closed with a VGG (visual geometry group) network, receives a 64*64 color image and increases its size four times as much as it was to reach 256*256 dimensions. As a result, this network produces high-quality images. Figure 2 shows how this network works.

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

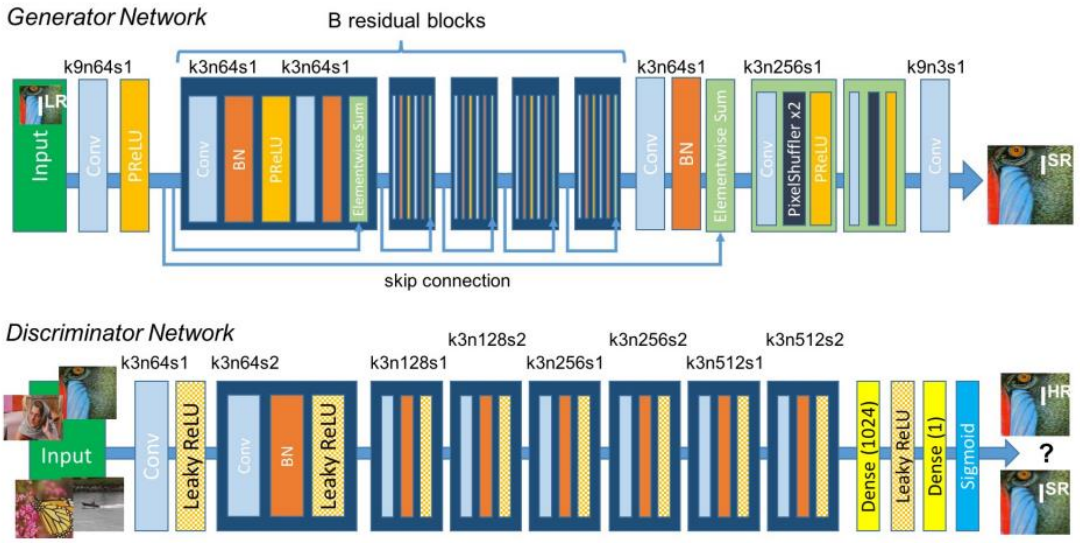

Figure 2. SRResNet [20].

To differentiate HR ground-truth pictures from super-resolution ones, a generator network and a discriminator network are utilized. Five residual blocks which change the image at the lower scale formed the generator network. This network also presents a method by ESPCN to reconstruct the super-resolution image without the need to complete intermediate pixel values [23]. Figure 2 provides a more detailed description of the method. Each generator's residual block provides two groups of the conventional (Conv-BN-ReLU) block, with a constant channel number of 46 and a kernel size of 3 for all blocks. A skip connection follows this. The Content Loss formula of SRResNet is:

$$l^{SR}_{VGG/i.j} = \frac{1}{W_{i,j}H_{i,j}}\sum_{x=1}^{W_{i,j}}\sum_{y=1}^{H_{i,j}}(\emptyset_{i,j}(I^{HR})_{x,y} - \emptyset_{i,j}(G_{\theta_G}(I^{LR}))_{x,y})^2 \quad (1)$$

Another network used in this model is RESNeXt, which is optimized for the SR issue. Figure 3 shows how this network works. The ResNeXt is used to solve problems related to a generative adversarial network for image super-resolution (SRGAN). The issues include complex computation, unstable networks, and the slow pace of learning. Using the ResNeXt network, a generator is created. This network decreases the computational complexity of the generator to a significantly lower amount (one-eighth) of the SRGAN. To solve the SRGAN's instability, the researchers created a discriminator using Wasserstein GAN (WGAN). The learning rate is boosted by omitting the normalization operation in the residual network. The experiments from the Res_WGAN indicated that the new model resulted in both subjective and objective positive assessments. These assessments were done using five datasets in comparison with the latest models [27].

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

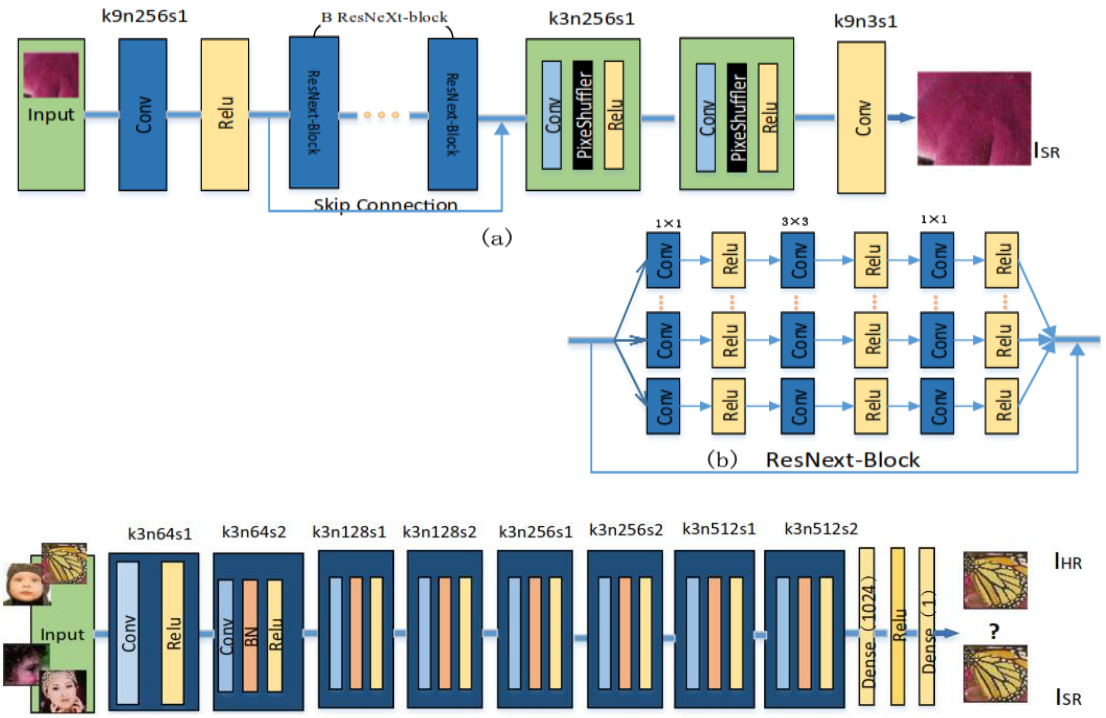

Figure 3. RESNeXt [24].

The loss function in RESNeXt is computed using this formula:

$$L_{Gen}^{SR} = \sum_{n=1}^{N} = D_{\theta_D}\left(G_{\theta_G}(I^{LR})\right) \qquad (2)$$

Another model which can be used to do the SR operation is SRCNN (super-resolution convolutional neural network). Figure 4 depicts the structure of the network of SRCNN and FSRCNN (fast super resolution convolutional neural network). There are some differences in the characteristics of FSRCNN and SRCNN. The first difference lies in the FSRCNN's use of the original low-resolution picture as input without bicubic interpolation. In addition, a deconvolution layer is added at the end to include upsampling. The second difference is that instead of using non-linear mapping steps in SRCNN, a three-step model including shrinking, mapping, and expanding steps was used. Third, FSRCNN uses filters with smaller sized and a deeper network structure. These developments enable FSRCNN to perform better than SRCNN and impose lower computational costs [28].

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

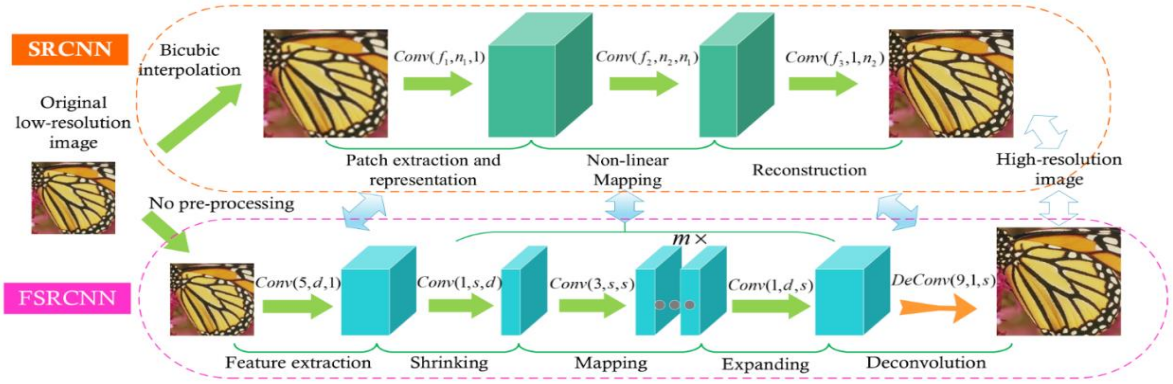

Figure 4. SRCNN [28].

The loss function in SRCNN is computed using this formula:

$$loss = \sum_{i=1}^{N} a y_i log\ (p_i) + (1-\alpha)(1-y_i)\log(1-p) \qquad (3)$$

The last network to be included in this model is efficient sub-pixel convolutional neural networks (ESCPN)[17]. SISR is used to examine an HR image ISR in a condition in which an LR image ILR is downscaled from the corresponding HR image IHR. The downsampling operation is straightforward: Creating ILR from IHR requires us to convolve IHR by employing a Gaussian filter. This can stimulate the camera's point spread function. Then, downsampling the image by a factor of r is done. Here, r is the upscaling ratio. Both ILR and IHR can have C color channels; therefore, they are characterized as real-valued tensors of size H × W × C and rH × rW × C. The SRCNN, which recovers from the upscale and interpolated version of ILR, is employed to solve the SISR problem. The recovery of ISR includes the use of a three-layer convolutional network. This novel network architecture avoided upscaling ILR before inputting it into the network. In the new architecture, we first employ a one-layer convolutional neural network to the LR image. Then, we apply a sub-pixel convolutional layer that upscales the LR feature maps to produce ISR [22]. Figure 5 shows how this network works.

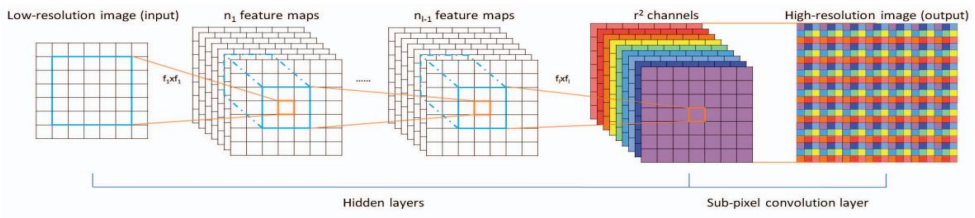

Figure 5. ESCPN

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

The process is that the LR image is first converted into HR image dimensions by interpolation methods. Therefore, the LR image data is set on a surface (or the equivalent of a manifold placed in an image-sized space). This procedure is different from the HR data procedure because the LR image has a certain distance from the HR images after interpolation (otherwise, the research on machine learning methods was worthless). The LR dimensions are clearly different from the procedure dimensions. HR will be less due to less detail. If Manifold Learning is performed in the LR procedure, the coordinates of the images on the procedure are obtained. The coordinates of the equivalent HR images on the HR procedure must be extracted from these coordinates. This can be done with feature learning methods. HR images must then be reconstructed from the learned Manifold HR coordinates [29, 30].

The classification problem is categorizing the new observations based on hypothesis H learned from the training data set. The hypothesis H represents a mapping of input data features to the appropriate target labels/classes. While learning the hypothesis H, the main objective is that it should approximate the true unknown function as close as possible to reduce the generalization error. Several applications of these classification algorithms range from medical diagnosis to remote sensing [31].

HR images play an essential role in different fields, including satellite imagery, aerial imagery, ultrasound imaging, medical imaging, traffic monitoring, security surveillance, ground-based remote sensing, astronomical observation, and biometric information detection. Due to hardware and physical limitations, producing high-quality images is costly and causes problems such as low signal-to-noise ratio (SNR) and prolonged image production time [32].

III. PROPOSED METHOD

In this paper, Ensemble deep learning has been employed to improve an image's resolution. In the experimental part of our study, we employed five pre-trained deep neural networks. We set their inputs as 64*64 pictures (images with small dimensions) and their outputs as 512*512 pictures (images with the original dimensions). Figure 6 depicts the overall process of the proposed method. After employing upsampling, the processed low-resolution (PLR) image is created. Then, the GAN networks (including LapSRN, SRResNet, RESNeXt, SRCNN, and ESCPN), fined-tuned by our datasets, is used to generate five super-resolution counterpart images. Finally, a Convolutional Neural Network (CNN) performs ensemble learning and achieves the final image.

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

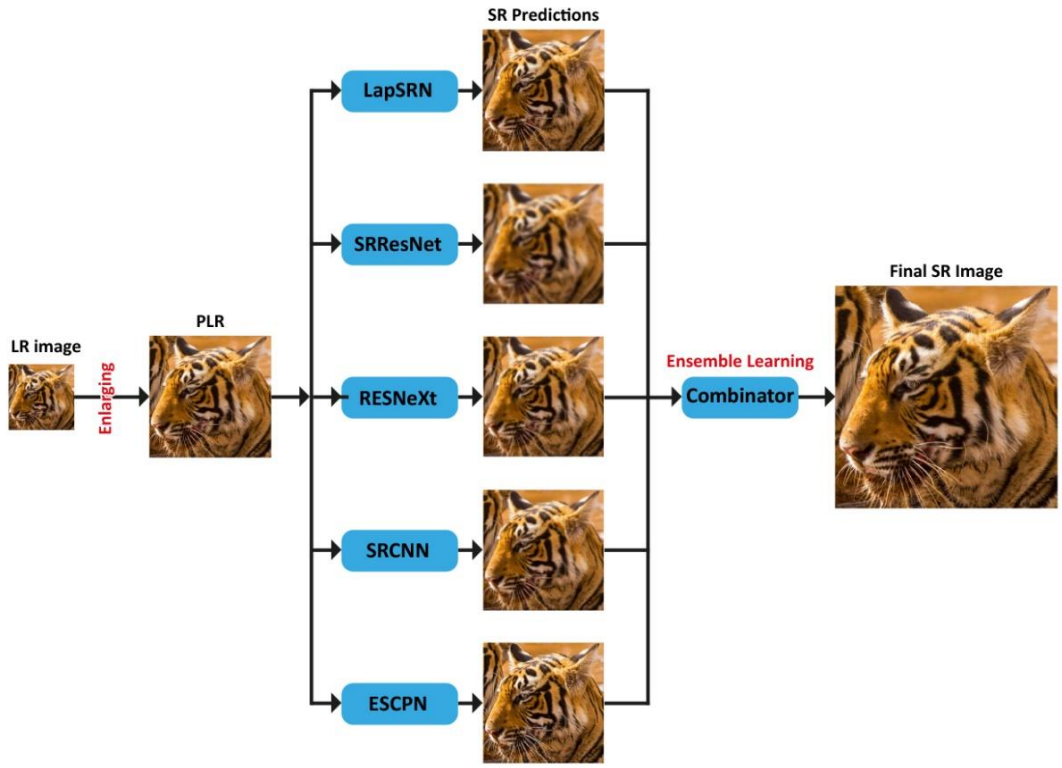

Figure 6. Workflow of the SR method used in this study

This study assumes that $I_{LR} \in \mathbb{R}^{p\times q}$ is a low-resolution image, and the image size is p × q. The corresponding HR image is $I_{HR} \in \mathbb{R}^{m\times n}$, and m×n is the size of the HR image.

$$I_{LR} = (I_{HR}) \qquad (4)$$

Here, the $I_{HR} \in \mathbb{R}^{m\times n} \rightarrow I_{LR} \in \mathbb{R}^{p\times q}$ reflects the process of blurring, which creates a low-resolution image from an HR one. Then, an approximate inverse function ($g \approx f^{-1}$) is applied:

$$g \approx f^{-1} : I_{HR} \approx g(I_{LR}) \quad (5)$$

Since it is impossible to identify this inverse function, one can employ different priors to reach the closest value to f^{-1}. The proposed method can be discussed in three main stages.

1) **Pre-processing**: The first step is down-sampling, in which the HR images were turned into the frequency domain using the 2D Fourier transform. The focal 25 percent of the points were kept in the k-space, and all peripheral data points were zeroed out. Using the 2D inverse Fourier transform, the low-resolution dataset was turned into an image. This process resulted in low-resolution images that were half of the original images in size; in other words, $p = m/2$

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

and $q = n/2$. The output of the previous stage is a processed low-resolution image (PLR) that was enlarged to the size of the intended HR image. It is a common step since the enlargement can improve the quality of images, and the neural networks can extract image features more easily. In addition, we employed different SR algorithms to establish the datasets.

2) **GAN models**: The generators of different GAN models were employed in this study to generate the HR images from LR images. The GAN models used in this paper are LapSRN, SRResNet, RESNeXt, SRCNN, and ESCPN.

3) **Ensemble learning**: Different data sources are used in ensemble learning to reach the most efficient results. The super-resolution images, which were the results of different GAN models, were integrated using ensemble learning to improve SR image quality. A CNN model is employed to integrate the results of the five GAN models. The results of the GAN models are concatenated before being inserted into the CNN. Images from different results were concatenated before entering the CNN for training, validation, and testing.

During the training process, the mean-absolute-error (MAE) is also calculated. The learning process employed in the ensemble process is as follows

$$I_{SR} = (\sum I_{SR,} , \Theta) \qquad (6)$$

where $ISR \in \mathbb{R}\ m \times n$ indicates ensemble learning outputs (our SR outcomes), ϕ: $ISR,i \in \mathbb{R}\ m \times n \rightarrow ISR \in \mathbb{R}\ m \times n$ suggests the process of ensemble learning, Θ represents the parameters of the CNN model.

IV.Experimental Results

The used datasets and the evaluation metrics are provided in this section.

4.1 Datasets

We used DIV2K and BSDS300 datasets in this research to assess the proposed method and compare it with state of the art in this realm. Table 1 shows the specifications of two used datasets.

Table 1: Datasets that have been used in this paper

Dataset	Training data	Validation data	Experimental data
DIV2K	800 images	100 data	100 images

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

(RGB images)	(with three low-quality images equivalent)	(with equivalent low-quality images.)	(with equivalent low-quality)
BSDS300	200 images	----------	100 images

1. DIV2K: This dataset contains RGB images, which are divided into three parts:

a) Training data: 800 images with three low-quality images equivalent to different quality reduction coefficients.

b) Validation data: includes 100 data with equivalent low-quality images.

c) Experimental data: Includes 100 images with equivalent low-quality images.

Images are reduced by the bicubic method as specified in the file name.

2. BSDS300: This dataset was originally used for segmentation, so the images in the dataset should be used to produce low-quality images. This dataset contains 200 training images and 100 experimental images.

4.2 Evaluation Metrics

Mean square error (MSE) for measuring image quality is the most common metric obtained by averaging the difference in brightness of each pixel of the reference image and the image being measured. The MSE in the formula below shows that $A_{i,j}$ is the original image, and $B_{i,j}$ is the reconstructed image.

$$\mathrm{MSE} = \frac{\sum_{j=1}^{N}(\sum_{i=1}^{M}(A_{i,j} - B_{i,j})^2)}{\mathrm{MN}} \quad (7)$$

Here, M is the number of rows in image A, and N is the number of columns in image B. In addition, the PSNR criterion, which is a function of MSE, is used in research:

$$\mathrm{PSNR} = 10\log_{10}\frac{(\text{maximum pixel value})^2}{\mathrm{MSE}} \quad (8)$$

These criteria were widely used due to their ease of use, calculation, and statistical support. Although MSE is highly studied theoretically and has good asymptotic properties, it is not visually satisfactory [23, 24]. In the last three decades, a lot of effort has been made to develop image evaluation methods. These efforts have been made in two main directions, one is the study of the characteristics of the human visual system (HVS) and the attempt to find equivalent computational criteria for its simulation [25], and the other is the geometric study

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

of the data in Input space and attempt to provide a criterion that is consistent with the geometric structure of the data [32, 29].

Most of the work done is centered on the first approach, and in the meantime, most efforts have been made to add a penalty to MSE to limit its free parameters [25]. From the first approach, the structural similarity index measure (SSIM) is presented in [26] according to the function of HVS. The performance of this criterion is illustrated in Figure 7. Although this criterion is designed based on the function of the human visual system, it approaches the second approach because it seeks to extract the structural similarity of the data.

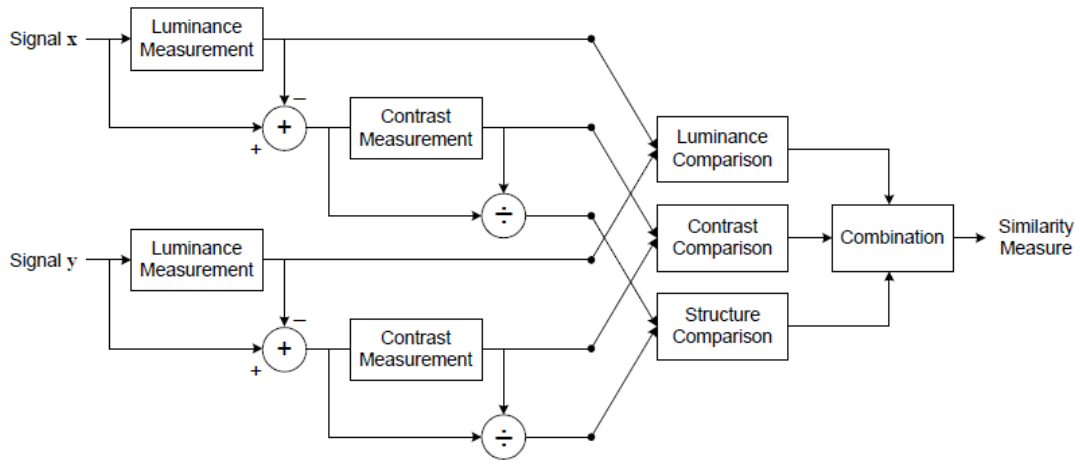

Figure 7. SSIM performance diagram [22]

The structure of an object in a scene is independent of its brightness. Therefore, to examine the structural information in an image, the effect of brightness must be separated. Because the brightness and contrast vary from scene to scene, local brightness and contrast are used to define the criterion. The system shown in Figure 7 divides the similarity criterion into three parts: brightness, contrast, and structure. First, the brightness is compared. Suppose x and y are two sub-images with non-negative values of one image. Then, the average of each sub-image is used to compare the brightness:

$$\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad (9)$$

The brightness comparison function l (x, y) is a function of μ_x and μ_y. By removing the mean from the sub-image, the sub-image is displayed on the following screen cloud:

$$\sum_{i=1}^{N} x_i = 0 \qquad (10)$$

The standard deviation is then used to estimate the contrast of the image:

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

$$\sigma_x = \left(\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)^2\right)^{\frac{1}{2}} \quad (11)$$

The contrast comparison function c (x, y) is a function of σ_x and σ_y.

The signal is normalized by dividing by the above standard deviation and comparing the structure of s (x, y) on this normalized signal. The above three components are combined to form the desired similarity criterion S (x, y). The definition of these three functions and how to combine them in the paper (Wang et al., 2004) is based on the following three conditions:

1. Symmetry: S (x, y) = S (y, x)
2. Boundary: S (x, y) ≤1
3. Having a maximum value: S (x, y) = 1 if and only if x = y

The function l (x, y) is defined as follows:

$$l(x,y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (12)$$

The constant C_1 is defined to prevent values close to zero of the denominator. This constant is considered as follows:

$$C_1 = (K_1 L)^2 \quad (13)$$

Where L is the image signal range (for example, for a grayscale image with 8 bits per pixel, this value is 255) and K_1 is a small value less than 1: K≪1. The function l precisely satisfies the three required conditions. In addition to these three conditions, this criterion is also compatible with the Weber rule or a light mask used to model light in the human visual system. The contrast comparison function also has a similar structure:

$$c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (14)$$

Where C_2 is defined as $C_2 = (K_2 L)^2$. This function is also compatible with the contrast mask used in HVS.

After the image of each of the sub-images below the space equivalent to the equation $\sum_{i=1}^{N} x_i = 0$ and normalizing it with standard deviation, to measure the similarity of the structure, we can find the correlation between the two Used images. Since the correlation

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

between $\frac{x-\mu_x}{\sigma_x}$ and $\frac{y-\mu_y}{\sigma_y}$ is equal to the Pearson correlation coefficient between x and y, the function can be defined as follows:

$$s(x,y)=\frac{\sigma_{xy}+C_3}{\sigma_x\sigma_y+C_3} \qquad (15)$$

Where σ_xy is defined as follows:

$$\sigma_{xy}=\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\mu_x)(y_i-\mu_y) \qquad (16)$$

The general criterion defined in [22] combines the above three parts as follows and then specifies the default value for the specified parameters α, β, and γ, and the criterion used in the article is achieved:

$$SSIM(x,y)=l(x,y)^{\alpha}.c(x,y)^{\beta}.s(x,y)^{\gamma} \qquad (17)$$

These three parameters in the article are all considered equal to 1, and in addition, the value $C_3=\frac{C_2}{2}$ is set:

$$SSIM(x,y)=\frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)} \qquad (18)$$

In general, PSNR is used to evaluate the similarity of the gray value and SSIM is used to show the structural similarity. Using these two criteria, different methods are compared together. Other algorithms can be used to calculate MSE. When images are corrupted by noise, l_1 software can not eliminate dots, but preserves the texture of the image. In contrast, l_1 software has many problems in dealing with Gaussian noise distribution, and l_p software achieves better results in terms of visual impact and quantitative indicators. For example, in Figure 8, the soft l_p with $p = 1.3$, and in Figure 9, the soft $l_p$with $p = 1.5$, are provided. Also, it can be clearly seen in Figure 8 that l_1 and l_p are stronger than l_2 [29, 30].

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

Figure 8. Results of reconstruction of transferability of Lena image using (a) Bilinear interpolation and (b) soft l_2, (c) soft l_1, and (d) soft l_p with $p = 1.3$. As shown in Figure 8, image (a) has noise, and its edges are not obvious. The input parameters in (b), (c), and (d) determine the clarity value of the edges of these images; however, these parameters increase the complexity and operation time of the algorithm. In fact, there is a trade-off between time, complexity, and quality.

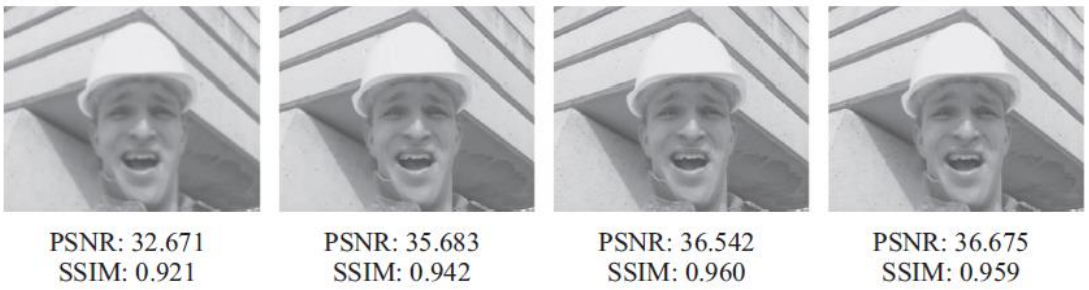

Figure 9. Results of reconstruction of the resolution of the image without noise using (a) Bilinear interpolation and (b) soft l2, (c) soft l1, and (d) soft lp with $p = 1.5$

V. Results

The results of the study are provided in this section. We applied the proposed algorithm to modify some images. Some of these result samples are presented in this section. Figures 10 and 11 provide the results of the implementation of the proposed method.

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

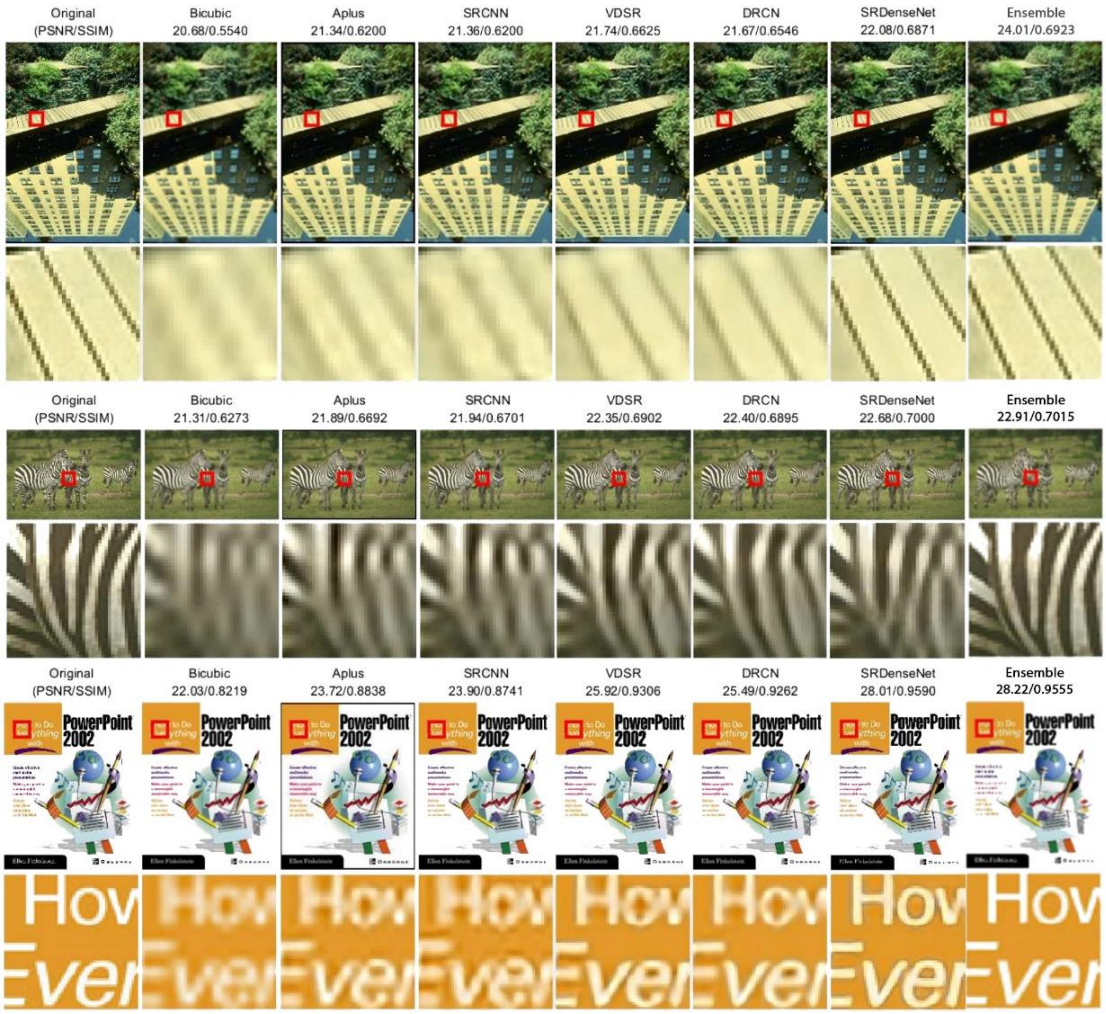

Figure 10. Results of our proposed method against other methods

Figure 10 provides the results of the comparison of reconstructed images obtained from different methods. As shown in Figure 10, severe distortions were detected in the reconstructed results which were formed using previous methods, but the proposed method proved to be more successful in reconstructing the texture pattern and minimizing the distortions (e.g. for the last image in figure 10 PSNR = 28.22 and SSIM = .9555).

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

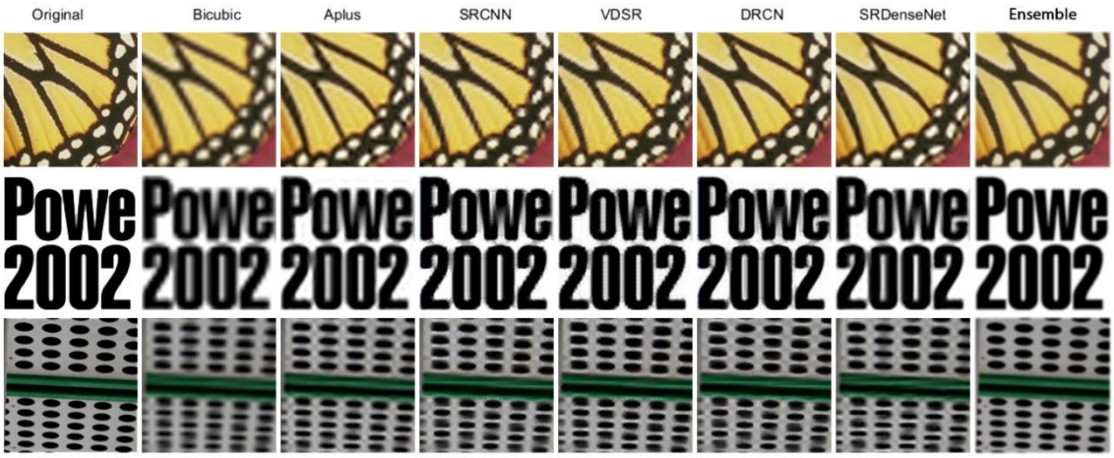

Figure 11. Results of our proposed method

The results show that the blurring effect was reduced, and the quality of colors was not significantly different from the ground truth. This minimal change might be due to the fact that the color means obtained from the employed networks were not used in the ensemble method, and we benefited from the optimum result extracted from the networks.

As the results show, the present model employed the extracted knowledge and shared it within different networks. The results of this procedure were significantly better than those of previous methods. Of course, the weaknesses in each network could adversely affect its performance, but using different networks can minimize these weaknesses and result in high-quality products.

Table 2. The comparison of the results of our proposed method with those of others using DIV2K and BSDS300 datasets

	nearest	**bicubic**	**Aplus**	**SRCNN**	**VDSR**	**DRCN**	**SRDenseNet**	**Proposed method**
DIV2K								
PSNR	26.36	28.33	30.17	30.33	31.52	30.76	32.05	**33.40**
SSIM	0.7542	0.8221	0.8617	0.872	0.8938	0.8784	0.9019	**0.9172**
MSE	22.21	23.11	22.70	21.594	22.55	21.25	20.47	**21.7**
BSDS300								

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

PSNR	24.46	26.29	26.48	26.85	27.92	28.16	27.19	**28.13**
SSIM	0.7100	0.7586	0.7861	0.7962	0.8174	0.8004	0.8084	**0.8497**
MSE	23.43	25.57	24.75	23.96	25.41	24.34	23.59	**22.05**

Instead of the comparison method (bicubic interpolation), we use the pre-defined VGG of a conditional comparison network. Extensive assessments in the benchmark dataset indicate that the proposed model provides better results than SRGAN. As shown in Table 2, the results of the quantitative appraisal criteria indicate that the proposed method was significantly more efficient than the previous methods in four (out of six) measures. In this method, in addition to reducing the quantity of data to train each section (Single Image) and using pre-trained networks to decrease the time required, we employed the ensemble method to improve the quality, time spent, and complexity.

VI. Conclusion and Future Work

This study was conducted to examine how ensemble learning models and SR could be integrated to achieve HR images out of LR ones. To this end, we proposed a deep convolution network within the framework of the Laplace pyramid for an extremely rapid and accurate resolution of a single image. Our model gradually predicts the remaining high-frequency residual frequency. We used DIV2K and BSD 300 datasets to fine-tune five pre-trained models. The results of this study showed an improvement of 0.1 dB in PSNR. Thus, it can be argued that our proposed model can be taken as a universal method to optimize the quality of SR methods that use machine-learning methods. Furthermore, our model presents a way to optimize the quality of the results of SR methods by using ensemble deep learning models. The proposed method can achieve PSNR = 33.4 and SSIM = 0.9172 in average on the DIV2K dataset and also PSNR = 28.13 and SSIM = 0.8497 on the BSDS300 dataset that proved to be more successful in reconstructing the texture pattern and minimizing the distortions.

Future work can be conducted to examine the effect of using a GAN network instead of a combinator on the quality of the results. We believe that using the statistical relationship is less accurate than a trained network in an image translation task, and the quality of the results does not improve significantly; however, a trained network can lead to better results. This change in the proposed method can be examined in other studies. In addition, future models can enable users to select their desired size to determine the quality.

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

REFERENCES

[1] Marivani, I., Tsiligianni, E., Cornelis, B., & Deligiannis, N. (2019, September). Multimodal image super-resolution via deep unfolding with side information. In 2019 27th European Signal Processing Conference (EUSIPCO) (pp. 1-5). IEEE.

[2] Wang, Z., Chen, J., & Hoi, S. C. (2020). Deep learning for image super-resolution: A survey. IEEE transactions on pattern analysis and machine intelligence, 43(10), 3365-3387.

[3] Haris, M., Watanabe, T., Fan, L., Widyanto, M. R., & Nobuhara, H. (2017). Super-resolution for UAV images via adaptive multiple sparse representation and its application to 3-D reconstruction. IEEE Transactions on Geoscience and Remote Sensing, 55(7), 4047-4058.

[4] Xiao, M., Zheng, S., Liu, C., Wang, Y., He, D., Ke, G., ... & Liu, T. Y. (2020, August). Invertible image rescaling. In European Conference on Computer Vision (pp. 126-144). Springer, Cham.

[5] 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.

[6] Fukami, K., Fukagata, K., & Taira, K. (2019). Super-resolution reconstruction of turbulent flows with machine learning. Journal of Fluid Mechanics, 870, 106-120.

[7] Wang, Z., Chen, J., & Hoi, S. C. (2020). Deep learning for image super-resolution: A survey. IEEE transactions on pattern analysis and machine intelligence, 43(10), 3365-3387.

[8] Seryasat, O. R., & Haddadnia, J. (2018). Evaluation of a new ensemble learning framework for mass classification in mammograms. Clinical breast cancer, 18(3), e407-e420.

[9] Rahmani Seryasat, O., Kor, I., Ghayoumi Zadeh, H., & Shams Taleghani, A. (2021). Predicting the number of comments on Facebook posts using an ensemble regression model. International Journal of Nonlinear Analysis and Applications, 12(Special Issue), 49-62.

[10] Ghayoumi Zadeh, H. , Fayazi, A. , Rahmani Seryasat, O. and Rabiee, H. (2022). A Bidirectional Long Short-Term Neural Network Model to Predict Air Pollutant Concentrations: A Case Study of Tehran, Iran. Transactions on Machine Intelligence, 5(2), 63-76. doi: 10.47176/TMI.2022.63

[11] Tong, T., Li, G., Liu, X., & Gao, Q. (2017). Image super-resolution using dense skip connections. In Proceedings of the IEEE international conference on computer vision (pp. 4799-4807).

[12] Yang, C. Y., Ma, C., & Yang, M. H. (2014, September). Single-image super-resolution: A benchmark. In European conference on computer vision (pp. 372-386). Springer, Cham.

[13] Tanno, R., Worrall, D. E., Ghosh, A., Kaden, E., Sotiropoulos, S. N., Criminisi, A., & Alexander, D. C. (2017, September). Bayesian image quality transfer with CNNs:

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

exploring uncertainty in dMRI super-resolution. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 611-619). Springer, Cham.

[14] Timofte, R., De Smet, V., & Van Gool, L. (2014, November). A+: Adjusted anchored neighborhood regression for fast super-resolution. In Asian conference on computer vision (pp. 111-126). Springer, Cham.

[15] Caballero, J., Ledig, C., Aitken, A., Acosta, A., Totz, J., Wang, Z., & Shi, W. (2017). Real-time video super-resolution with spatio-temporal networks and motion compensation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 4778-4787).

[16] Mei, Y., Fan, Y., & Zhou, Y. (2021). Image super-resolution with non-local sparse attention. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 3517-3526).

[17] Talab, M. A., Awang, S., & Najim, S. A. D. M. (2019, June). Super-low resolution face recognition using integrated efficient sub-pixel convolutional neural network (ESPCN) and convolutional neural network (CNN). In 2019 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS) (pp. 331-335). IEEE.

[18] Kawulok, M., Benecki, P., Piechaczek, S., Hrynczenko, K., Kostrzewa, D., & Nalepa, J. (2019). Deep learning for multiple-image super-resolution. IEEE Geoscience and Remote Sensing Letters, 17(6), 1062-1066.

[19] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. Advances in neural information processing systems, 27.

[20] Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., ... & Shi, W. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4681-4690).

[21] Lai, W. S., Huang, J. B., Ahuja, N., & Yang, M. H. (2017). Deep laplacian pyramid networks for fast and accurate super-resolution. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 624-632).

[22] Kim, J., Lee, J. K., & Lee, K. M. (2016). Accurate image super-resolution using very deep convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1646-1654).

[23] Denton, E. L., Chintala, S., & Fergus, R. (2015). Deep generative image models using a laplacian pyramid of adversarial networks. Advances in neural information processing systems, 28.

[24] Shi, W., Jiang, F., & Zhao, D. (2017, September). Single image super-resolution with dilated convolution based multi-scale information learning inception module. In 2017 IEEE International Conference on Image Processing (ICIP) (pp. 977-981). IEEE.

Power System Technology
ISSN:1000-3673

Received: 20-10-2022 *Revised: 10-06-2023* *Accepted: 15-12-2023*

[25] Ding, K., Ma, K., Wang, S., & Simoncelli, E. P. (2021). Comparison of full-reference image quality models for optimization of image processing systems. International Journal of Computer Vision, 129(4), 1258-1281.

[26] Wang, Z., & Bovik, A. C. (2009). Mean squared error: Love it or leave it? A new look at signal fidelity measures. IEEE signal processing magazine, 26(1), 98-117.

[27] Nan, F., Zeng, Q., Xing, Y., & Qian, Y. (2020). Single image super-resolution reconstruction based on the ResNeXt network. Multimedia Tools and Applications, 79(45), 34459-34470.

[28] Dong, C., Loy, C. C., & Tang, X. (2016, October). Accelerating the super-resolution convolutional neural network. In European conference on computer vision (pp. 391-407). Springer, Cham.

[29] Haris, M., Shakhnarovich, G., & Ukita, N. (2018). Deep back-projection networks for super-resolution. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1664-1673).

[30] Guan, J., Zhang, W., Gu, J., & Ren, H. (2015). No-reference blur assessment based on edge modeling. Journal of Visual Communication and Image Representation, 29, 1-7.

[31] Xue, W., Mou, X., Zhang, L., Bovik, A. C., & Feng, X. (2014). Blind image quality assessment using joint statistics of gradient magnitude and Laplacian features. IEEE Transactions on Image Processing, 23(11), 4850-4862.

[32] Yue, L., Shen, H., Li, J., Yuan, Q., Zhang, H., & Zhang, L. (2016). Image super-resolution: The techniques, applications, and future. Signal Processing, 128, 389-408.