Enhancing image quality using super-resolution residual network for small, blurry images

Djarot Hindarto[1,2], Mochammad Iwan Wahyuddin[2], Andrianingsih Andrianingsih[3],
Ratih Titi Komalasari[2], Endah Tri Esti Handayani[3], Mochamad Hariadi[1]

[1]Department of Electrical Engineering, Faculty of Intelligent Electrical and Informatics Technology, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[2]Department of Informatics, Faculty of Communication and Information Technology, Nasional of University, Jakarta, Indonesia
[3]Department of Information Systems, Faculty of Communication and Information Technology, Nasional of University, Jakarta, Indonesia

Article Info

ABSTRACT

In the background, when low-resolution images are utilized, image identification tasks are frequently hampered. By employing the residual network super-resolution framework, super-resolution techniques are used to enhance image quality, specifically in the detection and identification of small and blurry objects. Improving resolution, decreasing blur, and enhancing object detail are the main goals of the suggested approach. The novelty of this research resides in its application of the activation exponential linear unit (ELU) to the super-resolution residual network (SR-ResNet) framework, which has been demonstrated to enhance image sharpness. The experimental findings demonstrate a substantial enhancement in the quality of the images, as evidenced by the training data's structural similarity index (SSIM) of 0.9989 and peak signal-to-noise ratio (PSNR) of 91.8455. Furthermore, the validation data demonstrated SSIM 0.9990 and PSNR 92.5520. The results of this study indicate that the implementation of SR-ResNet significantly enhances the capability of the detection system to detect and classify diminutive and opaque entities precisely. The expected and projected enhancement in image quality significantly influences image processing, especially in situations where accuracy and object differentiation are vital.

Corresponding Author:

Mochamad Hariadi
Department of Electrical Engineering, Faculty of Intelligent Electrical and Informatics Technology
Institut Teknologi Sepuluh Nopember
Surabaya 60111, Indonesia
Email: mochar@te.its.ac.id

1. INTRODUCTION

Advancements in image processing have greatly enhanced the ability to recognize and analyze objects. However, the detection of small and indistinct objects must be improved. The low image resolution poses a constraint that impedes tasks such as security monitoring, medical applications, and autonomous technology. It becomes difficult to discern intricate details in blurry images. In order to tackle the issue at hand, it is necessary to go beyond the achievements of traditional image processing methods and instead focus on fostering innovation and implementing a methodical approach. Super-resolution (SR) provides a promising solution to enhance image resolution and overcome these limitations. Nevertheless, further investigation is required to examine the capacity of SR to identify minute and indistinct entities. The integration of SR with the super-resolution residual network (SR-ResNet) architecture has demonstrated the potential to improve

image recognition. The limited image resolution poses significant challenges in accurately identifying and distinguishing objects, resulting in profound implications in domains such as medicine, security, and autonomous technology. Hence, the quest for solutions to detect diminutive and indistinct images continues to be a vibrant field of study.

Computer vision and image processing assess image quality. Structural similarity index (SSIM) and peak signal-to-noise ratio (PSNR) are used. SSIM takes structure, contrast, and luminance into account in order to quantify the degree of structural similarity between two images [1]. This analysis is more closely aligned with human perception compared to methods that solely rely on individual pixels. The SSIM values span from -1 to 1, with a value of 1 indicating complete similarity. PSNR quantifies the relationship between the highest signal intensity and the presence of noise in the signal representation. A high PSNR value, measured in decibels, signifies excellent image or video quality following procedures such as compression or transmission. Both metrics are commonly employed in research pertaining to image compression algorithms, image restoration, and other computer vision applications. The utilization of both SSIM and PSNR enables a thorough assessment of image quality, taking into account both perceptual and technical factors. Ensuring the delivery of superior outcomes that align with user specifications is paramount in image processing applications.

This research focuses on the SSIM value aimed at measuring the effectiveness of image processing and reconstruction compared to the original or reference image. SSIM is used as the primary metric due to its sensitivity to structural changes recognized by human visual perception. This research also seeks to improve image processing techniques to achieve higher structural similarity with the original image. This is very important in image restoration and compression, as well as in security monitoring, medical applications, and autonomous technology. Experiments conducted on training data show notable improvements in image quality evaluation using the PSNR metric of 91.8455 and SSIM of 0.9990. These results show that the proposed approach enhances the system's ability to improve image resolution and effectively identify small and blurry objects in the training data environment. In addition, when conducting experiments at 4x scale, the use of identical evaluation metrics showed an increase in PSNR value to 92.5520 and SSIM to 0.9990 on the validation data. This validates that the suggested methodology not only achieves improved image quality but also consistently improves the system's ability to detect small objects at high image resolutions.

The primary contribution of this research is the successful integration of SR and object detection techniques. This integration has led to improvements in image resolution and increased accuracy in distinguishing small and blurry objects in images. The obtained results demonstrate that the integration of SR-ResNet yields substantial enhancements in image quality [2] and the system's capacity to identify objects that are challenging to discern in low-resolution images. These findings are significant because they can improve visual acuity in various image-processing applications that require accurate object detection, such as in the fields of medicine, security, and autonomous technology. Through enhancements in image resolution and enhanced detection capabilities for minute and indistinct objects, this method could pave the way for more advanced image processing capabilities, enabling more precise identification of objects.

Prior research, as enumerated in Table 1 [3]-[16], tends to prioritize the PSNR value over the SSIM value as the primary quality indicator. SSIM, an algorithm that assesses the similarity of two images by analysing their brightness, contrast, and structure, is deemed highly pertinent; however, it has yet to be exhaustively examined within the scope of this investigation. Hence, to achieve a more comprehensive understanding, this study incorporates an assessment of the SSIM value on the reconstructed and enhanced image, in addition to the PSNR value. Additionally, findings from prior research indicate that investigations employing a 2x image scale exhibit a more pronounced augmentation in both PSNR and SSIM. As a result, this study introduces an innovative approach by utilizing a fourfold increase in image scale. The impetus and motivation for this research stem from the identification of gaps in prior studies. By integrating PSNR and SSIM assessments and broadening the research scope to include images at a 4x scale, this study aims to offer a more comprehensive understanding of methods for enhancing image resolution and reconstruction quality. This method will likely generate a substantial contribution to our knowledge of image resolution enhancement technologies and close a gap in the literature.

The novelty of this research lies in the significant improvement in SSIM values achieved through the use of SR technology integrated with the SR-ResNet architecture. The main focus is on improving the system's ability to identify fine details and distinguish small objects in images, which is reflected explicitly in the improved SSIM values. This innovation marks an essential advancement in image processing, particularly in improving the structural similarity of images, which is vital in applications that require accurate and high-detail object recognition.

2.　METHOD
2.1. Proposed method

Figure 1 illustrates the proposed research methodology for enhancing image resolution. The resolution enhancement process is accomplished by employing the SR-ResNet model [17], [18], uses two datasets:

low-resolution and high-resolution. The ResNet architecture, which has been adopted, comprises fundamental layers, including convolution layers (conv), batch normalization, activation functions utilizing exponential linear unit (ELU) [19], and weight addition. The model trains the network to recognize patterns and characteristics in low-resolution using a dataset. Upon completion of the training process, the model is anticipated to possess the capability to comprehend the correlation between low-resolution and high-resolution images. During this stage, the SR-ResNet model endeavors to generate high-resolution images from low-resolution images. This approach entails an intricate procedure wherein the model acquires the ability to comprehend intricate elements, formations, and patterns in low-resolution images, subsequently endeavoring to reconstruct more distinct and elaborate high-resolution image.

Table 1. Research discussing super resolution

| Author	Algorithm	Dataset	Findings
[3]	DMN, RDN	MRI	DMN-based eliminates false details in natural image tasks and improves quantitative and qualitative. 0.62 DMN+RDN PNSR 22.78 SSIM.
[4]	SR-radial fluctuations (SRRF)	Microscopic images	A SRRF method reduces for large raw images, visualizes microtubule dynamics used in biomedicine.
[5]	SR-CNN, PCA, Image feature	Milling image	Enhanced micro milling double dictionary. SR method outperforms others. Mean PSNR: 43.62, 42.19, 43.20, 4
[6]	SRCNN	BSD100, Urban100	The MBFSR method, extracts deep features using multiple feature extraction modules to create a high-resolution feature map, PNSR: 22.73, SSIM: 0.7427
[7]	Two-branch crisscross generative adversarial network (TBCGAN)	BSD100, Urban100	Two-branch crisscross generative adversarial network (TBCGAN) achieves realistic and accurate SR images. PSNR: 24.20, SSIM: 0.7518
[8]	L2 gradient loss LPIPS and FID	Set5, Set4, Urban100	Proposed an image SR algorithm based second-order gradient (SG) loss. SSIM 24.98
[9]	SR–cryoCLEM	cryo-electron tomography	Effect of high-power laser illumination on cryosamples by apoferritin structure.
[10]	SRGAN, SRRESNET	Set14, BSD100	SSIM SRGAN = 0.7397, SSIM SR_RESNET = 0.8184
[11]	ESRGAN	ovine (Malta) brucellosis	SSIM ESRGAN = 0.7840
[12]	ISFMFMNe	COCO2017	SSIM ISFMFMNe = 0.8104
[13]	DRMSFFN	SET14, B100	DRMSFFN AVG SSIM = 0.83
[14]	RCAN-SG	SET4, SET5, Urban100	Efficiency comparisons of upscale 4X RCAN-SG SSIM 0.8175
[15]	DBAGAN	SRLP dataset	DBAGAN, SSIM 0.824
[16]	USRKDN	Set5, Set14, BSD100	USRKDN SSIM = 0.7990
Author	SR-RESNET	Div2K	Proposed an image SR-RESNET algorithm LR=0.000001, Optimisation Adam, activation ELU. SSIM = 0.9989, PSNR = 91.8455
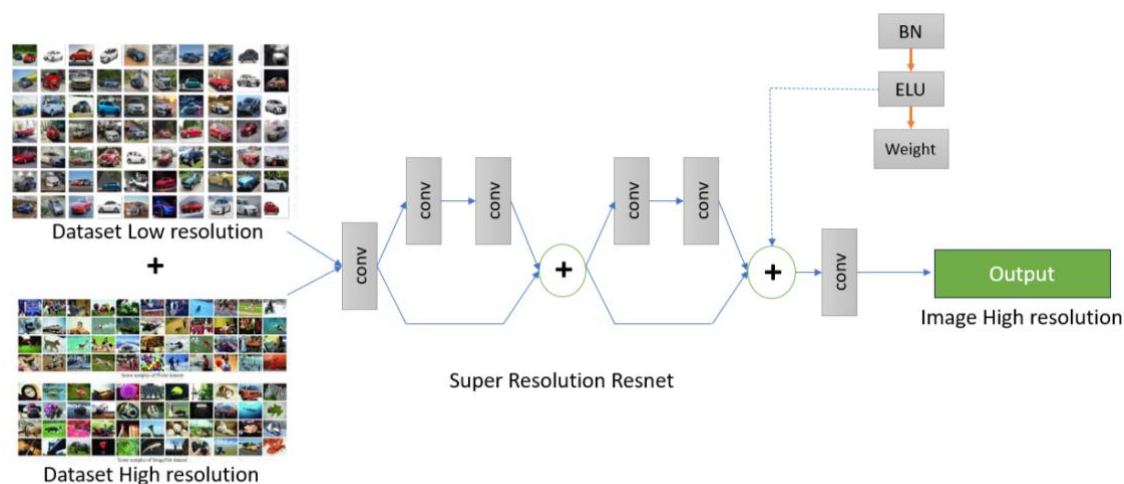
Figure 1. Proposed research method

The model uses filters to extract essential features from the imagery that is not high-quality during convolution. The ELU activation function aids the network in learning more intricate information [20]. Batch normalization is responsible for standardizing the output of the preceding layer, thereby accelerating the training procedure and enhancing the stability of the network. This process is iterated by incorporating weights

to fine-tune the network's weights, allowing them to gradually capture the correlation between low- and high-resolution. The outcome of this procedure is anticipated to be an enhanced high-resolution image derived from a low-resolution image. Figure 1 illustrates the iterative process of training the network using a low-resolution dataset, optimizing parameters, and testing the model to generate an image that is sharper and more closely resembles the high-resolution image as a whole. This method provides a robust approach to enhance the quality of low-resolution images, resulting in higher-resolution images that closely resemble the original image.

Table 2 explains input for the Res-Net algorithm used in SR-ResNet is an RGB image with 256×256 pixels. After the image is convolution with 64 filters using a 9×9 kernel size, the leaky rectified linear unit (LReLU) activation is applied to the convolution results. The output of the input image's convolution is the SR variable, which needs to be initialized first. Each iteration entailed two steps: first, the SR was subjected to the residual block process; second, the sum was computed between the SR and the residual block. This process was reiterated twice. Afterward, SR underwent convolution with 64 filters, 9×9 kernels, and the same padding. The process was repeated with 64 filters, 3×3 kernels, and the same padding. After that, the output of the convolution is superimposed onto the original picture. We rerun the same loop, this time applying convolution with 64 filters, 3×3 kernels, and the "same" padding to SR. Right after that, the SR undergoes a max pooling operation, which is followed by an up-sampling operation. The last step is to perform back convolution on SR using a 3×3 kernel, 64 filters, and the "same" padding as before. Finally, an image with three color channels created by ELU a during the convolution stage, the model applies filters to the image to extract significant features from the low-resolution image. The ELU activation function aids the network in learning more intricate information. Batch normalization is responsible for standardizing the output of the preceding layer, thereby accelerating the training procedure, and enhancing the stability of the network. This process is iterated by incorporating weights to fine-tune the network's weights, enabling them to progressively capture the correlation between low-resolution and high-resolution images. The outcome of this procedure is anticipated to be an enhanced high-resolution image derived from a low-resolution image. Figure 1 illustrates the iterative process of training the network using a low-resolution dataset, optimizing parameters, and testing the model to generate an image that is sharper and more closely resembles the high-resolution image. This method provides a robust approach to enhance the quality of low-resolution images, resulting in higher-resolution images that closely resemble the original image.

Table 2. Algorithm SR-ResNet

SR-ResNet
Input: image size 256×256, RGB
Operation:
image convolution, with filter 64, kernel 9×9, padding 'same'
ELU activation in convolution results
Define Model:
Initialize the SR variable, the result of convolution of the input image
Loop 2 times:
a. Residual Block to SR.
b. the sum between SR and Residual Block.
c. 64 convolution filters, 9×9 kernels, and 'same' padding on SR.
d. 64 filter convolution, 3×3 kernel, and 'same' padding on SR.
e. the summation between the SR convolution results and the input image.
Loop 1 time:
a. 64 filter convolution, 3×3 kernel, and 'same' padding on SR.
b. Max Pooling on SR.
c. up sampling operation on SR.
d. convolution with 64 filters, 3×3 kernels, and 'same' padding on SR.
Output: 3 color channel image using ELU activation and 'same' padding.

2.2. Dataset

Table 3 is a dataset containing numerical data, text, images, and other data for analysis and modeling [21]. Train, evaluate, and test machine learning models using training, validation, and test datasets. Dataset for image and video restoration (Div2K) is a collection of data intended for evaluation and development purposes in the field of image restoration [22]. The dataset is widely recognized for its exceptional quality and extensive assortment of images. Div2K comprises a collection of 1,760 low-resolution and high-resolution images captured under diverse conditions, such as low-light conditions, with a range of textures and levels of detail. Every image in this dataset has a high resolution, with the majority having a resolution of 2K (hence the dataset's name, Div2K). Due to its exceptionally high-resolution, it is a preferred option among researchers who are primarily concerned with image resolution enhancement or restoration techniques. Div2K empowers

users to conduct a comprehensive assessment of image processing algorithms and methodologies, with a particular focus on image restoration and SR techniques. By utilizing the variety of images supplied by Div2K, users can evaluate the efficacy of their algorithms across a spectrum of realistic scenarios. Div2K has emerged as a highly valuable dataset in the field of image processing research, particularly in the domains of image restoration and resolution enhancement. Its extensive collection of high-quality images has significantly contributed to the advancement of image processing technology.

Table 3. Dataset Div2K composition

Dataset	Number of datasets	Literature
Training high-resolution	685	[23]
Training low-resolution	685	[24]
Validation high-resolution	170	[25]
Validation low-resolution	170	[26]
Testing low-resolution	50	[27]

There are several low-resolution image datasets available for free on the internet. Here are some links that can be used to download low resolution image datasets: Div2k-dataset - A dataset consisting of low-resolution and high-resolution images used for object detection and segmentation tasks. In this research, the author collects datasets by mixing datasets from several image providers for free, so that the image datasets become diverse, so that later you can find out the strengths and weaknesses if the dataset is processed using the SR method. This research uses the SR-ResNet.

2.3. Super-resolution

SR is a technology to increase the resolution of images or videos from low resolution to high resolution by utilizing image processing techniques and machine learning [28]. SR technology is very useful when high resolution images or videos are needed [29], but only low-resolution images or videos are available. The super resolution approach can be done in two ways: traditional methods and deep learning methods. Traditional methods usually use interpolation techniques or signal processing to increase the resolution of images or videos. The deep learning method uses a neural network model to make predictions on the missing parts of an image or video and produce a higher resolution image or video [30]. There are several techniques used in super resolution, including,

- Interpolation: an interpolation technique [31] is used to increase the resolution of an image by adding pixels between the pixels already in the image. However, interpolation techniques often produce images or videos with poor quality, especially if the magnification is very large.
- Upscaling: an upscaling or upscaling technique is a technique used to increase the resolution of an image or video by doubling the pixels that are already in the image or video [32]. Upscaling techniques usually produce better quality images or videos than interpolation techniques [33].

SR is a technique for enhancing image quality by improving resolution and visual sharpness. This technology is applied in various fields, including photography, video streaming, medical image processing, and the movie industry. In photography, SR allows images taken at a low resolution to be enhanced to become more vivid and detailed. In video streaming, this technology helps to improve video quality by reducing blur and increasing sharpness. In medical image processing, SR is used to enhance medical images, enabling more accurate diagnosis. In the movie industry, SR is essential for improving the resolution of images or videos so that critical visual details can be seen more sharply. They are providing a more immersive and satisfying viewing experience.

Upscaling is a technique that enhances the resolution of an image or video by adding additional pixels. This method is beneficial for improving the visual appearance of low-resolution media [34]. Various algorithms, such as nearest neighbor, bilinear, bicubic, and lanczos, are employed in the upscaling process to improve the details of an image or video. Upscaling can increase resolution but may only sometimes enhance visual quality significantly and can appear poor if magnification is excessive. To address this issue, upscaling methods are frequently integrated with additional image processing techniques like denoising, deblurring, and deep learning to enhance the quality of images or videos by producing higher resolution and visually pleasing results. A more detailed mathematical formula for upscaling is as follows:

Let's say I_low is a low-resolution image, and I_high is the high-resolution image you want to produce. Upscaling is done by multiplying the I_low matrix by the T transformation matrix to get the I_high matrix:

$$I\_high\ (x, y) = T * I\_low\ (x, y) \tag{1}$$

Where (x, y) are the pixel coordinates in the image, and T is the transformation matrix. The transformation matrix T is found through different interpolation techniques, such as nearest neighbour, bilinear, bicubic, and lanczos.

2.4. Exponential linear unit

Neuronal networks are activated with the ELU [35]. ELUs [36], [37] unlike rectified linear unit (ReLUs), have negative values, pushing mean unit activations closer to zero, like batch normalization, with less computational complexity. A reduced bias shift effect brings the average gradient closer to the unit's natural gradient, promoting zero-speed learning. However, negative LReLUs and parametrized rectified linear unit (PReLUs) values do not guarantee noise-robust deactivation. Smaller inputs saturate ELUs to a negative value, reducing forward propagated variation and information. The ELU aids deep neural network learning and improves classification accuracy. ELUs mitigate the vanishing gradient problem like ReLUs, LReLUs, and PReLUs by identifying positive values. ELUs learn better than other activation functions. Formula ELU with $0 < \alpha$ is:

$$k(a) = \begin{cases} a, if\ a > 0 \\ \alpha(\exp(a) - 1), if\ a \leq 0 \end{cases} \tag{2}$$

$$k'(a) = \begin{cases} 1, if\ a > 0 \\ k(a) + \alpha, if\ a \leq 0 \end{cases} \tag{3}$$

3.   RESULTS AND DISCUSSION
3.1. Super-resolution residual network

This research aimed to conduct training experiments to enhance image resolution by employing the super resolution technique with a 4X scale. Upscaling is the process of improving the resolution of a low-resolution image to four times its original resolution. This experiment entails a training process of a SR model, which enables the conversion of low-resolution images into higher-resolution images. The objective of this experiment is to enhance the clarity, intricacy, and overall excellence of low-resolution images, transforming them into more transparent images with a higher resolution. By implementing a sequence of training procedures at a scale four times larger than the original, the super resolution model, specifically SR-ResNet, is educated to discern the underlying patterns and structures present in low-resolution images.

Consequently, it generates images with a resolution that is four times greater than the initial input image. This procedure entails iteratively optimizing the artificial neural network to enhance the model's capacity to estimate and restore intricate details that are diminished in low-resolution imagery. The utilization of 4X scaling in this experiment showcases the capability of the super resolution model to execute more extensive transformations, thereby posing additional difficulties in achieving precise reconstruction and enhancing the overall quality of the image. This research seeks to investigate the capabilities of the super resolution technique in addressing the problems associated with enhancing image resolution on a larger scale. It is anticipated that this technique will yield more satisfactory outcomes in the restoration of high-resolution images from low-resolution images. During the training process, the use of SR-ResNet to enhance image resolution from low to high has resulted in a collection of PSNR and SSIM values at each epoch. The outcomes of these ten epochs demonstrate substantial enhancements in the caliber of image reconstruction. In the beginning, during the first epoch, the PSNR and SSIM values were 21.2507 and 0.6619, respectively. At a low PSNR, the reconstructed image is significantly different from the original, indicating poor image reconstruction. Nevertheless, as the training iterations advanced, there was a steady enhancement in both evaluation metrics.

The PSNR graph in Figure 2 demonstrates a steady upward trend from the initial epoch to the fifth epoch, culminating in the highest value of 24.8211 in the fifth epoch. This signifies a notable enhancement in the resemblance difference between the original and reconstructed image. However, during the sixth epoch, while it continued to be at a relatively high level, there was a slight reduction in the PSNR value to 24.9382. Furthermore, the SSIM graph exhibits a comparable pattern. Commencing at a value of 0.6619 in the initial epoch, the SSIM progressively rises until it reaches its zenith in the tenth epoch, attaining a value of 0.8255. This signifies that the resemblance between the reconstructed image and the original structure, brightness, and contrast is approaching a state of near-perfect similarity.

However, it is essential to note that there are variations in some periods, exemplified by the seventh period, where the PSNR value decreased substantially to 22.1056. This shows that there may be fluctuations in the results due to the complex nature of the images or specific attributes of the training procedure. In Figure 3, graphs depicting the fluctuations in PSNR and SSIM values at each epoch offer a more in-depth understanding of the model performance. The image reconstruction improved from low resolution to high

resolution as the SR-ResNet training progressed, despite occasional fluctuations. This demonstrates the capacity of this technique to enhance image resolution significantly.
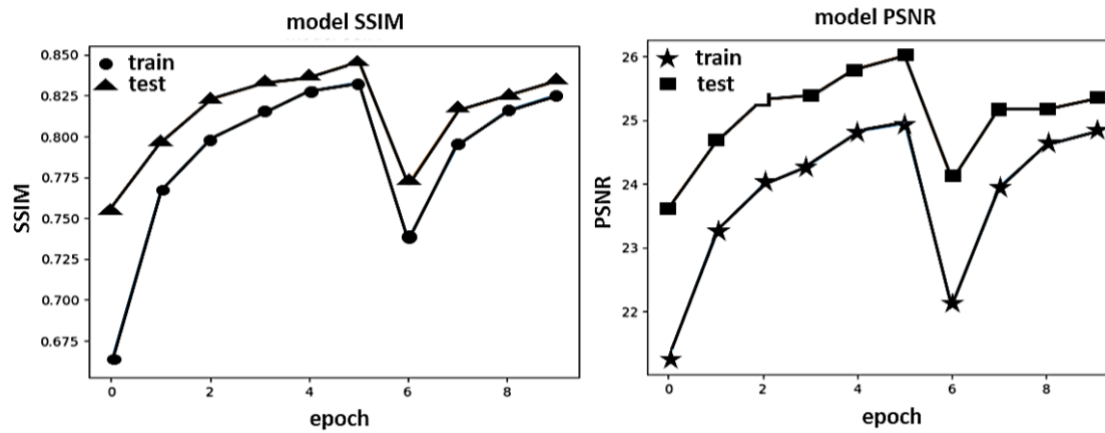

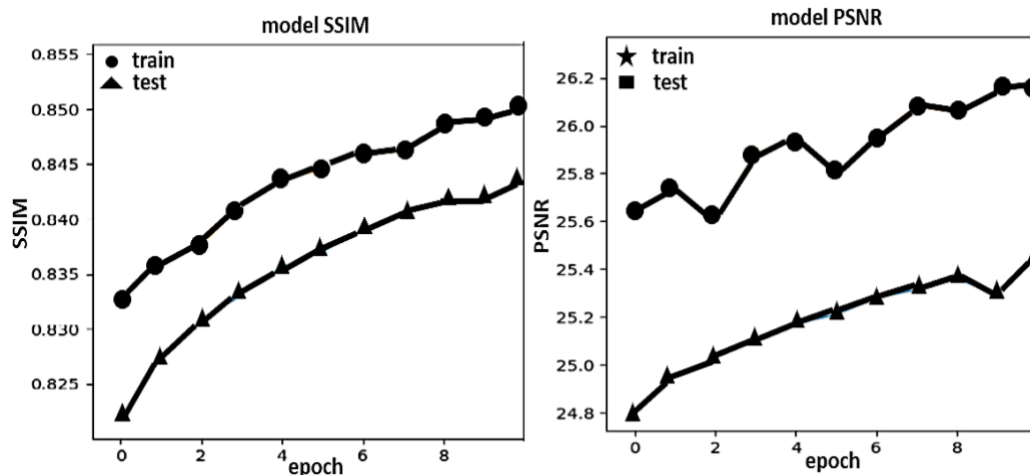Figure 2. Training model SSIM and PSNR

Figure 3. Training model SSIM and PSNR

For ten epochs, Adam's optimizer trained the model at 0.00001, improving image quality metrics like SSIM and PSNR. This improvement was seen in training and validation datasets. The SSIM, which measures structural similarity between images, increased from 0.8371 to 0.8474 in the training dataset and 0.8445 to 0.8543 in the validation dataset. This indicates that the model is becoming more effective at maintaining the visual integrity of the original images as the training progresses. The PSNR, a metric that quantifies the noise level in the reconstructed image in relation to the original image, demonstrates notable enhancements. In the training data, it increased from 25.2095 dB to 25.5490 dB, and in the validation data, it rose from 25.7978 dB to 26.2971 dB. This rise suggests a reduction in the image reconstruction error rate, leading to an improvement in the image reconstruction quality, approaching that of the original image.

The research proves that Adam's optimizer is helpful for training and validating picture reconstruction models. PSNR and SSIM were two measures of picture quality that the model consistently improved. On the training dataset, the model's improvement went from 91.8385 dB in the first epoch to 91.8455 dB in the tenth. On the validation dataset, it went from 92.5448 dB to 92.5520 dB, showing a steady improvement in reconstruction ability. SSIM, an essential measure of how healthy images are perceived, stayed high all the way through validation and training, rising slightly from 0.9990 in the first epoch to 0.9990 in the tenth. The model's consistent performance on hidden data suggests that it is balanced, as evidenced by its low variance and good convergence during training and validation. The modest gains in PSNR and SSIM, however, point to either untapped potential or a learning plateau with the provided data. The cautious conditions to prevent

overshooting gradient descent, which could slow down convergence, could be indicated by the low learning rate. With a stable SSIM value at a high threshold and a consistent improvement in PSNR, the final training results demonstrate excellent performance. Additional trials with varying learning rates are being contemplated to evaluate the model's efficacy without lowering the bar for generalizability or reconstruction quality. The explanation above can be seen in Figure 4.
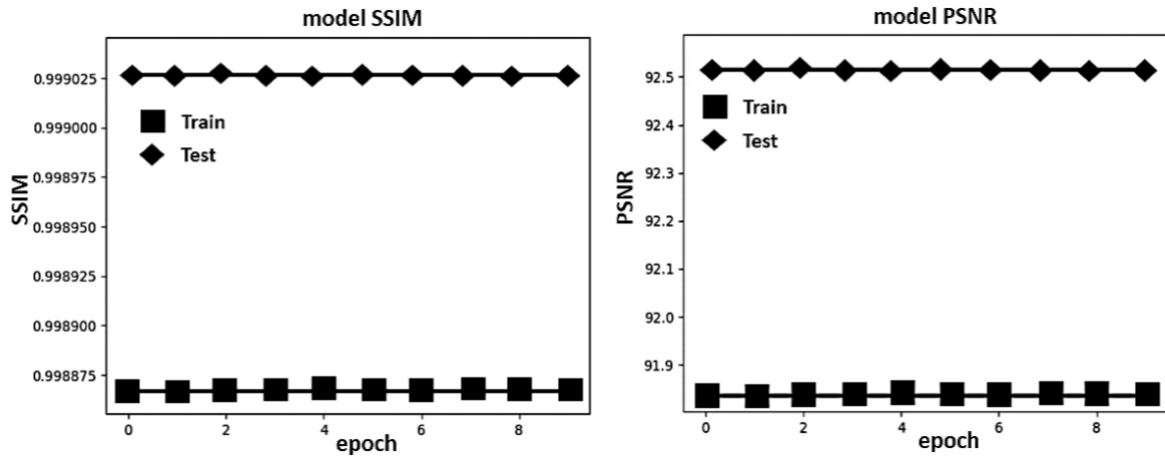

Figure 4. Training model SSIM and PSNR

Figure 5 shows tested the super resolution model, two metrics assessed image quality after reconstruction: SSIM and PSNR. We employed a variety of low-resolution images for this purpose. The first step in the testing process involves processing a red car image three times. The PSNR value increased consistently in both the first and second tests, reaching 22.14 dB and 23.35 dB, respectively, with an SSIM value of approximately 0.82. This proves that the model can enhance red car images with higher resolution. Nevertheless, despite a rise in PSNR to 23.35 dB in the third test using a blue car image, SSIM decreased. When an image's structure is altered, the reconstructed and original versions will look different from a structural perspective. A drop in PSNR to 23.35 dB and an increase in SSIM to 0.83 were observed in the most recent test on the white car image. Although the PSNR has decreased, the structural similarity between the original and reconstructed images has significantly improved.
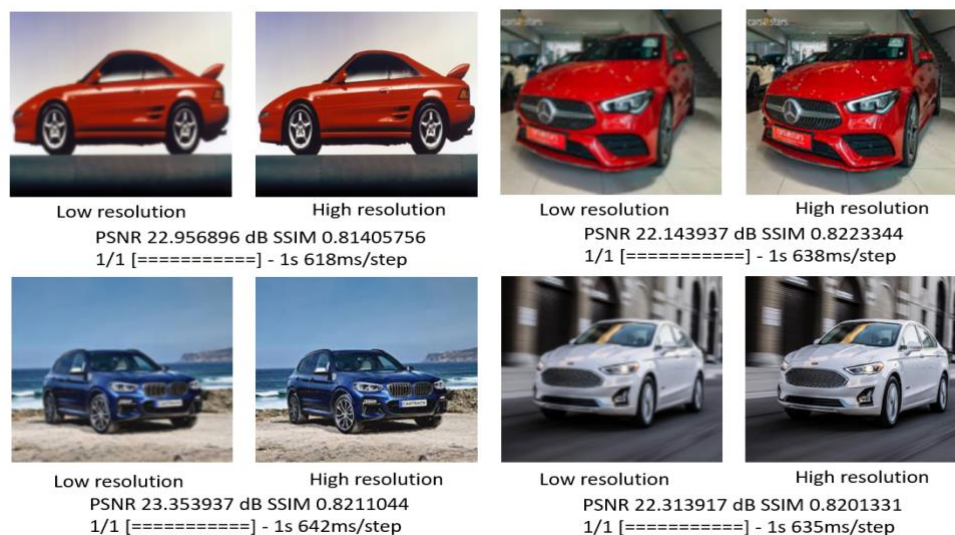

Figure 5. Testing low resolution to high resolution

The results show that the SR model's performance varies with the image's characteristics. Increasing PSNR does not necessarily ensure better structural similarity despite PSNR's widespread use as a critical

metric. To evaluate changes in reconstructed images, it is essential to have a better grasp of structural similarities; SSIM evaluation offers this. Because it paints a more complete picture of structural similarity—which may differ from PSNR values-SSIM should be prioritized as an evaluation metric. Regardless of the PSNR value, the key to comprehending the potential structural changes in the reconstructed image lies in the application of SSIM in SR evaluation. To guarantee consistently improved image quality, it is crucial to evaluate SR performance holistically. This research demonstrates that when SSIM is used in conjunction with PSNR, a more comprehensive and accurate picture of the reconstructed image's quality can be obtained. In this comprehensive analysis, the structural similarity measured by SSIM values is also considered so that the rise in PSNR does not solely dictate the increase in picture resolution. This lends credence to the idea that developing super resolution techniques necessitates a thorough evaluation to attain more consistent and satisfying outcomes.

3.2. Discussion, limitations and future work

In this research experiment, a Core I7 laptop featuring 16 GB of RAM and a 1 TB hard drive disk was used to carry out the SR-ResNet experiment. Opencv, tqdm, tensorflow, and numpy libraries, in addition to the python programming language in operating system Windows 10. Necessary library packages used in computational mathematics and data science include numpy, tqdm, tensorflow, and opencv. Constantly, twelve hours are devoted to training. The SR-ResNet model is generated through this procedure.

3.2.1. Dataset analysis

Div2K is a dataset that includes five primary subsets, one for each of the three main types of data: training, validation, and testing. The goal of using 685 high- and low-resolution images pairs in the training data is to help the artificial neural network grasp the connection between the two types of graphics. As a result, this aids the training process for transforming low-resolution images into high-resolution ones that are both clear and detailed. resolution for the sake of fair evaluation of the model's performance and to guarantee that it can generalize to new data as well as the training data, verification like this is crucial. At the same time, 50 low-resolution images make up the testing portion of this dataset; they are utilized to assess the model's robustness on novel data. Researchers can test the efficacy of restoring images of high resolution from low-resolution ones. The Div2K dataset offers enough variety in terms of both composition and subset size to conduct comprehensive and trustworthy training, validation, and testing of algorithms and models for the task of solving the challenge of re-creating high-resolution images from low-resolution ones.

3.2.2. Process training super-resolution residual network

The results of the training process using the SR-ResNet for ten epochs or iterations are shown in Tables 4 to 6. Every row in the table represents a training epoch and a model performance metric. The "epoch" column displays the precise epoch number during the training process. Additionally, multiple columns encompass evaluation metrics, including "loss," which represents the loss function value computed by the model during each epoch. This loss function measures the model's predictive accuracy in determining the target value.

Additionally, there is a column labeled "Acc," which indicates the accuracy of the model at each epoch. Accuracy measures the level of precision exhibited by the model in its predictions of targets. Following that, there is the PSNR column, which serves as a metric for evaluating the quality of image reconstruction, specifically in the context of high-level image restoration. PSNR quantifies the level of similarity between the reconstructed image and the original image in terms of both clarity and accuracy. A column called SSIM measures the similarity of the reconstructed image's structure or texture to the original. The column labeled "val" represents the validation process and displays the same metrics as loss, acc, PSNR, and SSIM, which are measured on validation data. Validation data is data that is not utilized during the training process to prevent overfitting. In general, this table presents a summary of how the model's performance changes over time as it undergoes training. As time passes, it is evident that the loss value decreases, while accuracy, PSNR, and SSIM show an upward trend in both training and validation data. This suggests the model can reconstruct high-resolution from low-resolution ones better. Despite occasional fluctuations in loss values during specific periods, the evaluation metric values on the validation data consistently demonstrated improvement, suggesting the model's strong ability to generalize.

3.2.3. Limitations

The SR-ResNet method, 4X image scale experiments, and ELU activation in this research have several drawbacks. With exploring lower scales like 2X or 3X, 4X experiments are unlimited. Experiments at a specific scale may lose important insights into SR-ResNet method performance and generalization at lower resolutions. Consider multiple scales in experiments to better understand how the method adapts to resolution changes.

This limitation may limit SR-ResNet's applicability in contexts with different resolution requirements. In addition, ELU activation, while helpful in vanishing gradient problems, can limit you if misused, especially in situations where other activation methods are better.

3.2.4. Future work

Given the limitations of the current study, extending the experiments to different resolution scales would help determine the SR-ResNet method's efficacy. In addition, investigating various activation function options may help better understand the complexity of this model. Although this research has provided a solid foundation, there is an urgent need to conduct a more comprehensive follow-up study. The main focus of future research should be on analyzing the broader resolution context and parameter variations in SR-ResNet. By addressing these limitations, it is expected that the SR-ResNet method can be further optimized for various SR applications. This follow up research could improve image resolution in photography, film, and medical image processing.

Table 4. Process training with optimizer=Adam (learning_rate=0.0001)

Epoch	Loss	Acc	PSNR	SSIM	Val loss	Val Acc	Val PSNR	Val SSIM
1	0.0094	0.7109	21.2507	0.6619	0.0053	0.7660	23.6326	0.7566
2	0.0059	0.7762	23.2355	0.7667	0.0043	0.8075	24.6491	0.7964
3	0.0050	0.8032	23.9899	0.7983	0.0037	0.8462	25.3208	0.8221
4	0.0047	0.8074	24.2967	0.8130	0.0037	0.8575	25.3848	0.8317
5	0.0042	0.8316	24.8211	0.8277	0.0034	0.8600	25.8173	0.8359
6	0.0041	0.8344	24.9382	0.8331	0.0032	0.8653	26.0199	0.8460
7	0.0193	0.7401	22.1056	0.7372	0.0047	0.7834	24.0669	0.7709
8	0.0049	0.8021	23.9617	0.7938	0.0038	0.8255	25.1737	0.8155
9	0.0043	0.8232	24.6017	0.8159	0.0037	0.8112	25.1761	0.8244
10	0.0042	0.8338	24.8172	0.8255	0.0036	0.8287	25.3313	0.8334

Table 5. Process training with optimizer=Adam (learning_rate=0.00001)

Epoch	Loss	Acc	PSNR	SSIM	Val loss	Val Acc	Val PSNR	Val SSIM
1	0.0039	0.8515	25.2095	0.8371	0.0034	0.8243	25.7978	0.8445
2	0.0039	0.8545	25.2795	0.8389	0.0033	0.8703	25.9431	0.8460
3	0.0038	0.8558	25.3142	0.8406	0.0032	0.8705	26.0831	0.8462
4	0.0038	0.8579	25.3580	0.8415	0.0032	0.8363	26.0554	0.8486
5	0.0038	0.8511	25.2991	0.8417	0.0032	0.8606	26.1499	0.8491
6	0.0037	0.8624	25.4504	0.8434	0.0032	0.8637	26.1724	0.8503
7	0.0037	0.8630	25.4675	0.8448	0.0031	0.8776	26.2104	0.8514
8	0.0037	0.8617	25.4825	0.8457	0.0031	0.8735	26.2445	0.8527
9	0.0037	0.8572	25.4916	0.8463	0.0031	0.8763	26.2723	0.8528
10	0.0037	0.8631	25.5490	0.8474	0.0031	0.8688	26.2971	0.8543

Table 6. Process training with optimizer=Adam (learning_rate=0.000001)

Epoch	Loss	Acc	PSNR	SSIM	Val loss	Val Acc	Val PSNR	Val SSIM
1	0.0036	0.8694	91.8385	0.9989	0.0031	0.8746	92.5448	0.9990
2	0.0036	0.8692	91.8393	0.9989	0.0031	0.8744	92.5459	0.9990
3	0.0036	0.8692	91.8400	0.9989	0.0031	0.8757	92.5464	0.9990
4	0.0036	0.8694	91.8409	0.9989	0.0031	0.8745	92.5474	0.9990
5	0.0036	0.8696	91.8417	0.9989	0.0031	0.8746	92.5485	0.9990
6	0.0036	0.8692	91.8424	0.9989	0.0031	0.8750	92.5492	0.9990
7	0.0036	0.8693	91.8432	0.9989	0.0031	0.8744	92.5495	0.9990
8	0.0036	0.8694	91.8440	0.9989	0.0031	0.8747	92.5505	0.9990
9	0.0036	0.8695	91.8447	0.9989	0.0031	0.8752	92.5513	0.9990
10	0.0036	0.8694	91.8455	0.9989	0.0031	0.8752	92.5520	0.9990

4.　CONCLUSION

The SR-ResNet model has successfully improved image resolution by applying SR techniques at a 4X scale. It effectively identifies patterns and structures in low-resolution images, resulting in images with four times higher resolution than the originals. However, challenges remain in achieving accurate reconstruction and improving image quality. The model showed significant improvement in image quality during training, with peak values of PSNR and SSIM at the fifth epoch. However, results varied over the training period, suggesting that image characteristics and certain training elements can impact the model's performance. Evaluation of SSIM is essential for a more comprehensive assessment of image quality. The study highlights the need for a comprehensive approach when developing SR techniques. Future research should explore the

use of the SRThis research has successfully demonstrated that the SR-ResNet model is capable of significantly improving the resolution of images through the application of SR techniques at a scale of 4 times. The model shows a unique ability to identify patterns and structures in low-resolution images and produce four times higher quality output. This represents a significant advance in efforts to improve image resolution. It lays the groundwork for broad practical applications in visual quality enhancement. In this study, the main findings are the significant increase in PSNR and SSIM values achieved by the SR-ResNet model, trial 1, learning rate 0,0001, ADAM optimizing, 5th epoch, PSNR=24,9382, SSIM=0,8331. Experiment 2, learning rate 0.00001, ADAM optimization, tenth epoch, PSNR=25.5490, SSIM=0.8474. Experiment 3, learning rate 0.000001, ADAM optimization, tenth epoch, PSNR=91.8455, SSIM=0.9989. This indicates that the model has great potential to produce images with much better quality than the original. However, the variability of the results over the training period indicates the significant influence of image characteristics and training parameters on the performance of the model. In this context, SSIM evaluation becomes very important to obtain a more accurate assessment of image quality. The results of this research have significant implications for the theory and practice of information technology and image processing. By demonstrating the effectiveness of the SR technique implemented by the SR-ResNet model, this research paves the way for the development of more advanced image resolution enhancement methods. This not only enriches the theory in the field of image processing but also provides practical solutions for applications that require image quality enhancement, such as medicine, security surveillance, and entertainment. This research has limitations, especially in terms of the variation in results caused by image characteristics and training elements. This suggests the need for further research to optimize the model by taking these factors into account. Future research can explore the applicability of the SR-ResNet model on a larger scale and different image types and seek the optimal balance between PSNR and SSIM improvement. A deeper investigation of the influence of image characteristics on model performance is expected to provide a deeper and more comprehensive understanding of SR techniques. This research provides new and essential insights into the potential and challenges of developing SR techniques. It provides a window of opportunity for future studies that push the boundaries of understanding and provide valuable solutions in the domains of computer vision and data processing.-ResNet model with different scaling scales and image types, optimize the balance between PSNR and SSIM enhancement, and investigate the influence of image characteristics on model performance. This will contribute to more advanced and versatile SR techniques for various applications.

ACKNOWLEDGMENTS

REFERENCES

[1] I. I. -Azcarate, R. D. Acemel, J. J. Tena, I. Maeso, J. L. G. -Skarmeta, and D. P. Devos, "4Cin: A computational pipeline for 3D genome modeling and virtual Hi-C analyses from 4C data," *PLoS Computational Biology*, vol. 14, no. 3, 2018, doi: 10.1371/journal.pcbi.1006030.

[2] Z. Lu, X. Jiang, and A. Kot, "Deep coupled ResNet for low-resolution face recognition," *IEEE Signal Processing Letters*, vol. 25, no. 4, pp. 526–530, 2018, doi: 10.1109/LSP.2018.2810121.

[3] X. Liu, H. Guo, H. Liu, and J. Li, "Domain migration representation learning for blind magnetic resonance image super-resolution," *Biomedical Signal Processing and Control*, vol. 86, 2023, doi: 10.1016/j.bspc.2023.105357.

[4] J. Chen *et al.*, "Deep-learning accelerated super-resolution radial fluctuations (SRRF) enables real-time live cell imaging," *Optics and Lasers in Engineering*, vol. 172, 2024, doi: 10.1016/j.optlaseng.2023.107840.

[5] S. Li, Z. Ling, and K. Zhu, "Image super resolution by double dictionary learning and its application to tool wear monitoring in micro milling," *Mechanical Systems and Signal Processing*, vol. 206, 2024, doi: 10.1016/j.ymssp.2023.110917.

[6] D. Li, S. Yang, X. Wang, Y. Qin, and H. Zhang, "Multi-branch-feature fusion super-resolution network," *Digital Signal Processing: A Review Journal*, vol. 145, 2024, doi: 10.1016/j.dsp.2023.104332.

[7] Q. Yang, Y. Liu, and J. Yang, "Two-branch crisscross network for realistic and accurate image super-resolution," *Displays*, vol. 80, 2023, doi: 10.1016/j.displa.2023.102549.

[8] S. Lin, C. Zhang, and Y. Yang, "A pluggable single-image super-resolution algorithm based on second-order gradient loss," *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, vol. 3, no. 4, 2023, doi: 10.1016/j.tbench.2023.100148.

[9] M. G. F. Last, W. E. M. Noteborn, L. M. Voortman, and T. H. Sharp, "Super-resolution fluorescence imaging of cryosamples does not limit achievable resolution in cryoEM," *Journal of Structural Biology*, vol. 215, no. 4, 2023, doi: 10.1016/j.jsb.2023.108040.

[10] C. Ledig *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017, pp. 105–114, 2017, doi: 10.1109/CVPR.2017.19.

[11] X. Wang, "An epidemiologic study on brucellosis in the vicinity of hohhot in China," *American Journal of Biomedical and Life Sciences*, vol. 6, no. 3, 2018, doi: 10.11648/j.ajbls.20180603.15.

[12] L. Fu, H. Jiang, H. Wu, S. Yan, J. Wang, and D. Wang, "Image super-resolution reconstruction based on instance spatial feature modulation and feedback mechanism," *Applied Intelligence*, vol. 53, no. 1, pp. 601–615, 2023, doi: 10.1007/s10489-022-03625-x.

[13] F. Liu, X. Yang, and B. De Baets, "A deep recursive multi-scale feature fusion network for image super-resolution," *Journal of Visual Communication and Image Representation*, vol. 90, 2023, doi: 10.1016/j.jvcir.2022.103730.

[14] W. Ying, T. Dong, and J. Fan, "An efficient multi-scale learning method for image super-resolution networks," *Neural Networks*, vol. 169, pp. 120–133, 2024, doi: 10.1016/j.neunet.2023.10.015.

[15] S. Pan, S. B. Chen, and B. Luo, "A super-resolution-based license plate recognition method for remote surveillance," *Journal of Visual Communication and Image Representation*, vol. 94, 2023, doi: 10.1016/j.jvcir.2023.103844.

[16] N. Yuan, B. Sun, and X. Zheng, "Unsupervised real image super-resolution via knowledge distillation network," *Computer Vision and Image Understanding*, vol. 234, 2023, doi: 10.1016/j.cviu.2023.103736.

[17] M. Duan *et al.*, "Learning a deep ResNet for SAR image super-resolution," *2021 SAR in Big Data Era, BIGSARDATA 2021 - Proceedings*, 2021, doi: 10.1109/BIGSARDATA53212.2021.9574228.

[18] W. S. Jeon and S. Y. Rhee, "Single image super resolution using residual learning," *2019 International Conference on Fuzzy Theory and Its Applications, iFUZZY 2019*, pp. 310–313, 2019, doi: 10.1109/iFUZZY46984.2019.9066214.

[19] D. Kim, J. Kim, and J. Kim, "Elastic exponential linear units for convolutional neural networks," *Neurocomputing*, vol. 406, pp. 253–266, 2020, doi: 10.1016/j.neucom.2020.03.051.

[20] Y. Li, C. Fan, Y. Li, Q. Wu, and Y. Ming, "Improving deep neural network with multiple parametric exponential linear units," *Neurocomputing*, vol. 301, pp. 11–24, 2018, doi: 10.1016/j.neucom.2018.01.084.

[21] D. Hindarto and A. Djajadi, "Android-manifest extraction and labeling method for malware compilation and dataset creation," *International Journal of Electrical and Computer Engineering*, vol. 13, no. 6, pp. 6568–6577, 2023, doi: 10.11591/ijece.v13i6.pp6568-6577.

[22] Y. Wan, M. Shao, Y. Cheng, D. Meng, and W. Zuo, "Progressive convolutional transformer for image restoration," *Engineering Applications of Artificial Intelligence*, vol. 125, 2023, doi: 10.1016/j.engappai.2023.106755.

[23] S. Gao *et al.*, "Implicit diffusion models for continuous super-resolution," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10021–10030, 2023, doi: 10.1109/cvpr52729.2023.00966.

[24] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1132–1140, 2017, doi: 10.1109/CVPRW.2017.151.

[25] S. H. Park, Y. S. Moon, and N. I. Cho, "Perception-oriented single image super-resolution using optimal objective estimation," in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1725–1735, 2023, doi: 10.1109/cvpr52729.2023.00172.

[26] A. Lugmayr, M. Danelljan, L. Van Gool, and R. Timofte, "SRFlow: learning the super-resolution space with normalizing flow," in *Computer Vision – ECCV 2020*, pp. 715–732, 2020, doi: 10.1007/978-3-030-58558-7_42.

[27] J.-E. Yao, L.-Y. Tsao, Y.-C. Lo, R. Tseng, C.-C. Chang, and C.-Y. Lee, "Local implicit normalizing flow for arbitrary-scale image super-resolution," in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1776–1785, 2023, doi: 10.1109/cvpr52729.2023.00177.

[28] J. Wu, Z. He, and L. Zhuo, "Video satellite imagery super-resolution via a deep residual network," *International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 2019, pp. 2762–2765, 2019, doi: 10.1109/IGARSS.2019.8900265.

[29] K. Karthik, S. S. Kamath, and S. U. Kamath, "Automatic quality enhancement of medical diagnostic scans with deep neural image super-resolution models," *2020 IEEE 15th International Conference on Industrial and Information Systems, ICIIS 2020 - Proceedings*, pp. 162–167, 2020, doi: 10.1109/ICIIS51140.2020.9342715.

[30] M. S. Habeeb, B. Aydin, A. Ahmadzadeh, M. Georgoulis, and R. A. Angryk, "Solar line-of-sight magnetograms super-resolution using deep neural networks," *Proceedings - 2020 IEEE International Conference on Big Data, Big Data 2020*, pp. 4586–4593, 2020, doi: 10.1109/BigData50022.2020.9378480.

[31] S. H. Nguyen and D. H. Phan, "Selective element domain interpolation technique for assumed rotations and shear strains in polygonal finite element thick/thin plate analysis," *Thin-Walled Structures*, vol. 186, 2023, doi: 10.1016/j.tws.2023.110677.

[32] W. Tan, N. Qin, Y. Zhang, H. McGrath, M. Fortin, and J. Li, "A rapid high-resolution multi-sensory urban flood mapping framework via DEM upscaling," *Remote Sensing of Environment*, vol. 301, 2024, doi: 10.1016/j.rse.2023.113956.

[33] T. Hunter, S. Hulsoff, and A. Sitaram, "SuperAdjoint: super-resolution neural networks in adjoint-based output error estimation," in *XI International Conference on Adaptive Modeling and Simulation (ADMOS 2023)*, pp. 1-8, 2023, doi: 10.23967/admos.2023.058.

[34] J. Panda and S. Meher, "An improved image interpolation technique using OLA e-spline," *Egyptian Informatics Journal*, vol. 23, no. 2, pp. 159–172, 2022, doi: 10.1016/j.eij.2021.10.002.

[35] K. Zhang, X. Yang, J. Zang, and Z. Li, "FeLU: a fractional exponential linear unit," *Proceedings of the 33rd Chinese Control and Decision Conference, CCDC 2021*, pp. 3812–3817, 2021, doi: 10.1109/CCDC52312.2021.9601925.

[36] B. Grelsson and M. Felsberg, "Improved learning in convolutional neural networks with shifted exponential linear units (ShELUs)," *2018 24th International Conference on Pattern Recognition (ICPR)*, Beijing, China, 2018, pp. 517-522, doi: 10.1109/ICPR.2018.8545104.

[37] H. Shen, Z. Wang, J. Zhang, and M. Zhang, "L-Net: A lightweight convolutional neural network for devices with low computing power," *Information Sciences*, vol. 660, 2024, doi: 10.1016/j.ins.2024.120131.

BIOGRAPHIES OF AUTHORS

Djarot Hindarto 🆔 📇 SC ⭕ received the B.Eng. degree in computer engineering from Sepuluh Nopember Institute of Technology (ITS), Indonesia, in 1994 and the Master of Information Technology Pradita University, in 2022, respectively. Currently, he is a lecture at the Department of Information Technology, University Nasional Indonesia. His research interests include security, artificial intelligence, deep learning, machine learning, data mining, game technology, internet of things, and blockchain. He can be contacted at email: djarot.hindarto@civitas.unas.ac.id.

Mohammad Iwan Wahyuddin received the B.Eng. degree in electrical engineering from, University of Northern Sumatra, Indonesia, Master of Gajah Mada University, Indonesia, in and Doctoral in electrical engineering from Indonesia University, Indonesia. Currently, he is a lecturer at the Department of Information Technology, University Nasional. His research interests include artificial intelligence, machine learning, and networking. He can be contacted at email: iwan.wahyuddin@civitas.unas.ac.id.

Andrianingsih Andrianingsih completed a doctoral degree in Information Technology at the Doctoral Program 2023. The Magister Program in Information Systems was completed between 2003 and 2006, while the bachelor's program in the same discipline commenced between 1998 and 2002 at Gunadarma University in Jakarta. She is affiliated with various international and national organizations, including the IEEE, the position of Head of the Information Systems Study Program. She has been working as a lecturer since 2003 at the National University. Conducting field study on the topics of information systems technology, geospatial analysis, business intelligence, communication data, and LoRa WAN. She can be contacted at email: andrianingsih@civitas.unas.ac.id.

Ratih Titi Komalasari is a lecturer and researcher at the Informatics Study Program, Faculty of Communication and Information Technology, National University of Jakarta since 2010. She has an undergraduate education background in Informatics, Gunadarma University, master's in information systems management at Gunadarma University and is currently pursuing doctoral studies in information systems at Diponegoro University. The teaching areas artificial intelligence, game programming, and data science. Meanwhile, the research field she is pursuing is data science, artificial intelligence, and geoinformatics. She can be contacted at email: ratih.titi@civitas.unas.ac.id.

Endah Tri Esti Handayani is a lecturer and researcher at the Information Systems Study Program, Faculty of Communication and Information Technology, National University of Jakarta since 2017. She has an undergraduate education background in Food Technology, Brawijaya University, Masters in Information Systems Management at Gunadarma University and is currently pursuing doctoral studies information systems at Diponegoro University. The teaching areas taught are statistics and probability, discrete mathematics, and management information systems. Meanwhile, the research field he is pursuing is data science. She can be contacted at email: endahtriesti@civitas.unas.ac.id.

Mochamad Hariadi is a lecturer and researcher at the Information Technology Department Computer Engineeering, Faculty of Electrical Engineering and Intelligent Informatics, Institut Teknologi Sepuluh Nopember, since 1996. He has an undergraduate education background in Bachelor Electrical Engineering (Institut Teknologi Sepuluh Nopember), Magister Information System (Tohoku University), Doctoral Computer Science and Mathematics (Tohoku University). He can be contacted at email: mochar@te.its.ac.id.