An Intelligent Deep Learning Architecture Using Multi-scale Residual Network Model for Image Interpolation

Diana Earshia V. * and Sumathi M.

Department of Electronics and Communication Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu 600119, India; Email: sumagopi206@gmail.com (S.M.)
*Correspondence: earshy@gmail.com (D.E.V.)

***Abstract*—Image interpolation techniques based on learning have been shown to be efficient in recent days, due to their promising results. Deep neural networks can considerably enhance the quality of image super-resolution, according to recent studies. Convolutional neural networks with deeper layers are commonly used in current research to improve the performance of image interpolation. As the network's depth grows, more issues with training arise. This research intends to implement an advanced deep learning mechanism called Deep Multi-Scaled Residual Network (DMResNet) for effective image interpolation. A network cannot be substantially improved by merely increasing the depth of the network. New training strategies are required for improving the accuracy of interpolated images. By using the proposed framework, the Low Resolution (LR) images are reconstructed to the High Resolution (HR) images with low computational burden and time complexity. In order to dynamically discover the image features at multiple scales, convolution kernels of various sizes based on the residual blocks have been utilized in this work. In the meantime, the multi-scaled residual architecture is formulated to allow these characteristics to interact with one another for obtaining the most accurate image data. The interpolation performance and image reconstruction efficiency of the proposed model have been validated by using a variety of measures such as PSNR, SSIM, RMSE, Run time analysis, and FSIM. Popular datasets IAPR TC-12, DIV 2K, and CVDS are used for validating the proposed model. This model outperforms the state-of-art interpolation techniques in its performance, by yielding an increase of 8% in PSNR, 6% in SSIM, 1.2% in FSIM, and a decrease of 38.79% in RMSE, 5.875 times in run time analysis.**

***Keywords*—image interpolation, high resolution, low resolution, Deep Multi-scaled Residual Network (DMResNet), feature extraction and reconstruction**

I. Introduction

Super-resolution imaging is a crucial step in the image synthesis process that raises the quality of the visual image. Super-resolution imaging is used extensively in many different industries, including those of medicine, imagery from satellites, and astronomy. In recent years, super-resolution imaging is being performed using learning techniques, particularly deep learning. This approach's methods have good quality levels; however, they don't have the right performance constraints. Individual requirements for communication and information processing have gradually expanded over the past several years as a result of the rise of the Internet and the quick development of information technology; image processing is a significant component of the processing of information. In addition, the technology known as image Super-Resolution (SR) is crucial for image processing. The idea is to use information processing technologies to transmit a number of Low-Resolution (LR) images to the final High-Resolution (HR) image. The conventional approach entails interpolating the LR tiny image to the necessary size, then using the reconstruction technique to generate an HR image. The super-resolution reconstructed image can be used broadly because it includes sharp details and an extensive amount of knowledge.

A well-known and challenging research topic in image processing is image interpolation, which is a form of super-resolution [1, 2]. It has been applied to a variety of tasks, including monitoring, scaling of images and videos, radiology, and much more. An individual image's resolution is intended to be increased using image interpolation techniques, a special sort of image super-resolution, in order to match the requirement for a high resolution [3]. Image interpolation is the process of creating a corresponding High Resolution (HR) image from a Low Resolution (LR) image that has been observed while preserving its sharp edges and detailed textures. Typically, the image interpolation methodologies are split into the following types [4]: learning based models, edge directed models, and polynomial based models. Classical polynomial-based interpolation techniques generate a perpetual interpolation function using the known pixel values, which is then used to resample the function to fill in the missing pixels [5]. These techniques produce real-time performance and work very well smooth images. Yet, because the fixed non-adaptive filters are applied in a variety of local pixel patterns, their results frequently exhibit blurring, haze, and interference effects close to the edges [6]. As a result, these techniques are inappropriate

Manuscript received May 22, 2023; revised July 6, 2023; accepted July 28, 2023; published September 22, 2023.

for images with intricate edges and patterns. Instead of interpolating the missing pixels across the edge, the edge-directed interpolation algorithms strive for this goal. To estimate the edge's direction in an HR image from its LR counterpart is typically the initial and most important step in these techniques [7]. Nevertheless, it frequently fails, particularly in areas with intricate surfaces and patterns.

In contrast to image super-resolution [8], image interpolation usually estimates the missing pixels when the LR images are simply down-sampled versions of the HR images with no pre-filtering, or distortion. Since numerous HR images can be produced from a single LR observation and multiple times of pixels must be inferred from the little data, this problem is poorly stated. Several interpolation methods [9, 10] have been developed in the literature during the last few decades. Due to its low computational cost and real-time speed, the interpolation approach is particularly well-liked and frequently utilized in real-world applications for real-time video processing. The polynomial-based interpolation methods, on the other hand, are only effective in areas of smooth images that often produce obtrusive artifacts near the edges, such as distortions, fading, halos, etc. [11]. This is primarily due to the interpolation's non-adaptive linear filter, which is unable to adjust to the various local pixel patterns. Their main concept is to interpolate along the edge orientation by using the edge information rather than interpolating from across edge. Edge-directed interpolation techniques may be separated into explicit methods and implicit methods [12, 13]. Some edge-directed interpolation techniques have been created to solve the issue. Their main concept is to interpolate along the edge orientation by using the edge information rather than interpolating from across edge. Edge-directed interpolation techniques may be separated into explicit methods and implicit methods. In the explicit approaches, the edge directions are explicitly determined before interpolation along the identified edge orientation [14]. Unfortunately, the interpolation quality is restricted by the edge orientation prediction accuracy.

The advancement of machine learning and deep learning has allowed learning-based image up-sampling techniques to advance rapidly over the past decade [15, 16]. These techniques' fundamental goal is to investigate the connection between the LR image patches and the associated HR image patches from either an independent image database or the original image. The system can produce images with more details can perform much better. Among other models, the learning-based interpolation techniques performs better, hence which are extensively used by the researchers in the literature works. Because of their excellent fitting capabilities, deep learning-based approaches model non—linearity using a lot of parameters [17]. However, it faces some significant challenges in terms of computational complexity, lot of training time, and difficulty in system design.

Therefore, the proposed work intends to develop a new machine learning based model for performing an image interpolation, which mainly focuses on the following objectives:

- To develop a new deep learning-based architecture for an effective image interpolation, the Deep Multi-scale Residual Network (DMResNet) algorithm is employed.
- To acquire the image features for reducing the computational complexity of reconstruction, a multi-scaled residual block has been constructed.
- To obtain the superior performance, a new weight updation model is incorporated for training the network.
- To assess the performance of the proposed interpolation scheme, the popular benchmarking datasets have been utilized along with the variety of metrics.

The remaining sections of this paper are divided into the following groups: The full literature review of the image interpolation techniques utilized in the previous works is presented in Section II, where the drawbacks and difficulties of the traditional techniques are also covered. The suggested machine learning-based image interpolation approach is briefly described in Section III along with its flow and descriptions. Moreover, Section IV employs a range of metrics to validate the outcomes of both the current and suggested interpolation processes. In Section V, the challenges and future scope of the article are summarized.

II. Literature Review

This section examines various image interpolation techniques that have traditionally been used to increase the resolution of LR to HR images. It also analyses the advantages and disadvantages of the methods employed in this study. These analysis are based on the simulation using Matlab software, run on Personal Computer with Intel i5 with 8GB RAM. Zheng *et al.* [3] developed a learning-based image interpolation algorithm that used a weighted direct nonlinear regression model to interpolate low-resolution images from high-resolution ones. The algorithm employed overlapping image patches and a refinement approach to improve the connection between LR and HR patches. The authors used a cascade framework to address the small training data set issue and avoid under- or over-training. Jakob *et al.* [18] developed an efficient sub-pixel Convolutional Neural Network (CNN) model for solving an image interpolation problem. In this work, a feed-forward neural network has been utilized to perform non-linear mapping for interpolation. Pandey *et al.* [19] utilized a CNN based image resolution method for enhancing the generation of high resolution images. It entails an end-to-end mathematical mapping between LR and HR images by non-linear feature extraction. In order to obtain better image smoothing, the standard CNN model is incorporated with the super-resolution imaging technique that effectively reconstructs the HR image. Huo, *et al.* [20] introduced a new NN algorithm with the integration of non-linear barycentric weight function for image interpolation. This framework includes the operations of image decomposition, clustering, non-linear mapping, and index mapping. The primary advantage of this work is, it effectively preserves the texture of image with minimal time cost.

Hou *et al.* [21] implemented an effective image resolution & reconstruction model based on the standard CNN algorithm. The input LR image is primarily smoothened, deblurred, and up-sampled in this research with the help of interpolation based super-resolution reconstruction techniques. The non-linear mapping is mainly performed to map the low dimensional and high dimensional features. Then, the reconstruction process is carried out based on the mean value of convolution kernel. Yet, the suggested framework is highly complex to implement, which could be the key problem of this work. *Garima et al.* [22] utilized a fractal interpolation based convolution methodology for the generation of high resolution images. Here, the light weight CNN architecture is deployed to improve the performance of image interpolation. Ismail *et al.* [23] implemented an Artificial Neuro-Fuzzy Inference System (ANFIS) based image resolution approach for interpolation. Typically, any LR image can be converted into a HR image using an Image Super Resolution (SR) techniques. Due to a number of factors, the most prevalent of which is the inability to afford a high resolution camera, LR images are used frequently in the real world applications. The learning-based approach has grown in popularity as a result of the advancement of machine learning techniques. First, linear mappings are specifically suggested to explain how LR and HR images relate to one another. In this research, a unique ANFIS interpolation-based technique for single frame image SR is given, which is especially useful when the training data is scarce. The training data is all presented at the beginning of the learning process, however the current approach executes ANFIS interpolation in a static manner. The Multilayer Perceptron (MLP) is the type of machine learning algorithm, which is used in this work for effectively performing an image interpolation. It can have a number of hidden layers in the middle in addition to input and output layers. In this model, a predefined activation function is used to processes the acquired result. The most important activation function known as, hyperbolic tangent function is computed in this model. Here, the weight and bias values are adjusted by the model for obtaining the required output based on the given input data.

Zhou *et al.* [24] introduced a dense convolutional auto-encoder model for an effective image interpolation. The authors created a brand-new framework based on multi-temporal term feature learning and several spatial scales. To execute non-linear mapping from the LR space to the HR space, features are taken from various time ranges and varied spatial resolutions. Wang *et al*. [25] developed a K-Means Principal Component Analysis (PCA) mechanism incorporated with the Gradient Descent (GD) model for an efficient image interpolation. Here, the image reconstruction is performed by using the local and global dictionaries. The artifacts caused by the feature discrepancy between the test image and the global training dataset are lessened by this local dictionary since it is more compatible with the source. Ismail *et al*. [26] performed an ANFIS based image interpolation to effectively map the LR into HR. Several low resolution, sub-pixel aligned images of an identical image are needed for multi-frame image SR techniques. Several learning-based SR approaches have been put forth as machine learning techniques advance, along with a variety of algorithms that can create learned mappings that mimic the intrinsic link between the LR and HR images. Hui *et al.* [27] designed a light weight multi-distillation model for the reconstruction of LR into a HR images. Step-by-step hierarchical features are extracted by the distillation unit, and the fusion module combines them based on the significance of potential features, which is determined by the suggested contrast-aware channel attention mechanism. Aydin *et al.* [28] deployed a new super-resolution model based on the Discrete Cosine Transformation (DCT) technique for image interpolation. The authors trained a deep neural network to convert low-resolution input images into high-resolution ones using the DCT frequency representation. Finally, the resulting frequency domain representation is converted back into the spatial domain, and unwanted effects are removed using a pre-trained artifact reduction model in the spatial domain (see Table I).

TABLE I. SUMMARY OF THE LITERATURE REVIEW

Ref.	Methods	Advantages	Disadvantages
Zheng *et al.* [3]	Weighted direct nonlinear regression model	Increased efficiency and reduced overfitting	Not suitable for handling complex image datasets and high time requirement.
Jakob *et al.* [18]	Sub-pixel CNN	Effective image interpolation and low time consumption	Increased error rate.
Pandey and Ghanekar [19]	CNN based image resolution	High accuracy	It is well-suited for large datasets, and high computational complexity.
Huo *et al.* [20]	Non-linear barycentric weighted Neural Network	Reduced time cost and better edge preservation	High processing complexity and overhead.
Hou *et al.* [21]	Effective image resolution & reconstruction based on CNN	It is better capable of handling both low and high dimensional features.	High computational and time complexities.
Ismail *et al.*[22]	Fractal interpolation based CNN model	Improved performance.	Difficult to understand the system model, and only suited for complex datasets.

III. PROPOSED METHOD

The field of image processing known as image SR is now extremely active. One or more LR images are used to calculate a HR image, where HR denotes greater features and LR denotes lesser features. Nevertheless, due to hardware device constraints and expensive cost, it occasionally collects only LR images. Hence, developing a precise and quick super-resolution technique is essential for improving image quality. Current super-resolution techniques can be generally categorized into a number of groups. Among others, a learning based image interpolation techniques are well-suited for effectively solving SR problem. Moreover, the nearest neighbor and bicubic are the widely preferred and standard methods for image interpolation. Since the learning based methods are more powerful and suitable for solving super relems. This kind of approach often begins with two sizable training data sets, one with low-resolution images and the other with high-resolution images. It can apply the estimated relation to a specific LR image to produce a HR result by learning the relationship between these LR and HR images. The main purpose of this work is to develop a novel and intelligent learning based image interpolation system for mapping LR to HR images. For this purpose, an advanced Deep Multi-Scaled Residual Network (DMResNet) mechanism is developed. Patches for learning mappings should function more effectively to improve the models' accuracy.

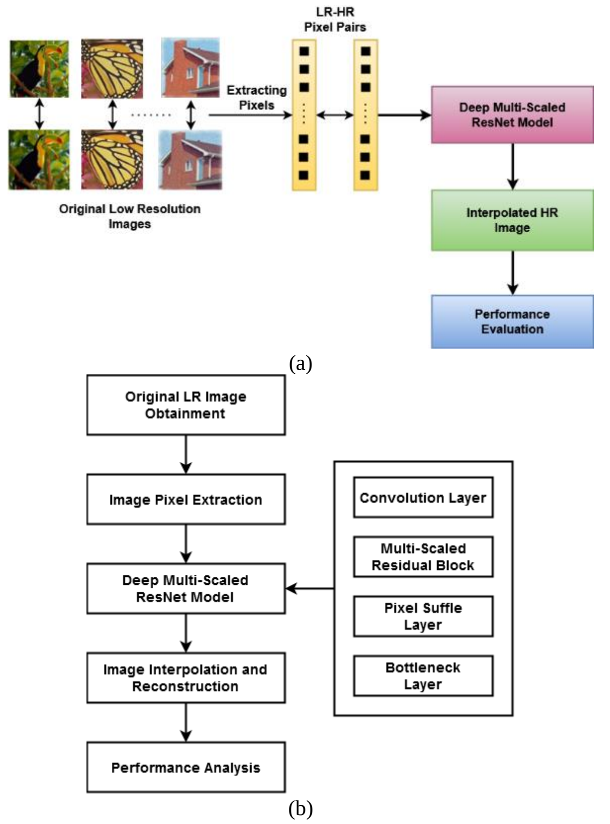


Figure 1. (a) Work flow of the proposed image interpolation framework. (b) Working stages of the proposed image interpolation system.

The work flow model of the proposed system is shown in Fig. 1(a) and the stages involved in the proposed system is illustrated in Fig. 1(b). In the proposed work, the new deep learning architecture model, called as, DMResNet is used as the feature extraction technique, which supports to obtain the multi-scale features from the low resolution original input images. During this feature extraction process, the low resolution image is taken as the input for processing, then the mapping function, loss function, and superior loss function are estimated with the use of multi-scaled residual blocks. It comprises the convolution and ReLU layers that are used to extract the image features at varying scales based on bypass operation. After that, these extracted features are fused to interpolate the high resolution image.

Bicubic interpolation, which is quick and effective in the smooth zone, is used to initially initialize the small LR images to the size of the HR images during training and interpolation. However, the bicubic will add artifacts to the edges and textures. Hence, it is essential to obtain an edge patch for training. Similar to this, the bicubic interpolation is applied for the other components and, just interpolate the edge patches to save time. Moreover, the edge patches on the bicubic images are determined with the help of statistical characteristics. The generated pixel pairs from LR-HR images serve as the training data set. It also makes the same assumption as the majority of other approaches already in use in the literature: there is enough training data to handle the image SR issues. Unfortunately, obtaining specific photographs or specific portions of images can be exceedingly challenging in many real-world circumstances. Due to the dearth of available image training data, the SR challenge remains a difficult one. The most recent deep learning-powered interpolation technique may offer a potential solution to this kind of issue because it has the ability to build new models in a given problem space using only a small amount of training data by utilizing well-trained models in the surrounding regions. The input and output images in image super resolution have a strong correlation. It is essential to completely use the input image's features and convey them to the network's endpoint for reconstruction. These traits, eventually fade away during transmission as network depth rises. This issue has led to a number of solutions, the most straightforward and effective of which is the skip connection. These techniques all aim to establish various connections between various layers. However, these techniques can't make full use of the input image's attributes and produce an excessive amount of redundant data that serves no useful function. Here, it is seen that as depth expands, the network's capacity for spatial expression gradually declines while its capacity for semantic expression steadily rises. Additionally, each MSRB's output has unique characteristics. The quality of the reconstructed images will therefore be directly impacted by how to utilize these hierarchical features to their best potential. This approach uses a straightforward hierarchical feature fusion model. The entire MSRB output is transferred to the network's end for restoration.

A. *Deep Multi-scaled Residual Network (DMResNet) Mode*

According to the recent literature works, it is analyzed that the deep learning based methods could highly improve the quality of super resolution. Typically, this kind of reconstruction is could be the one of the most challenging tasks, hence the learning based models are highly preferred for image interpolation. The conventional deep learning based architectures limit with the major problems of difficulty in image reproduce due to lack of network configuration, insufficient utilization of features, and low scalability. Therefore, the proposed work intends to utilize a novel deep learning model for improving image super resolution. The main purpose of this paper is to generate an interpolated HR image from the LR image with the use of deep learning algorithm, named as, DMResNet. In this algorithm [29, 30], LR original image (X_{LR}) is obtained as the input, which is efficiently reconstructed into a HR image (X_{HR}) using the proposed DMResNet model. To fully use the image features, we developed a novel DMResNet model in this study that outperforms the majority of state-of-the-art techniques. Here, the convolution kernels of various sizes are used based on the residual block that adaptively detects the image's characteristics at various scales. To obtain the most accurate image information, these features can interact with one another in the meantime; hecne it is referred to as multi-scale residual network.

At first, the bicubic operation is performed after obtaining the LR image (X_{LR}), and the image is converted into the YCbCr space with the training on Y channel. Here, the input image (X_{LR}) having C color channels is described with the tensor size of $\omega \times G \times C$. Then, the (X_{LR}) and (X_{HR}) are represented with $k\omega \times kG \times C$, where C is equal to 1 that indicates the Y channel, and k denotes the up scaling factor. Consequently, the end-to-end mapping function is computed for the reconstruction. For the given training $\{X_{LRi}, X_{HRi}\}_{i=1}^{N}$, the model has been constructed as shown in Eq. (1).

$$\delta = \arg\min_{\delta} \frac{1}{N} \sum_{i=1}^{N} Ls_{SR}(F_{\delta}(X_{LRi}), X_{HRi}) \qquad (1)$$

where, Ls_{SR} is the loss function, δ comprises the set of bias β and weight ω values of the network having size m as represented in the following form, as shown in Eq. (2).

$$\delta = \{\omega_1, \omega_2, \omega_3 \dots \omega_m, \beta_1, \beta_2, \beta_3 \dots \beta_m\} \qquad (2)$$

Moreover, the superior loss function like Mean Squared Error (MSE) and loss function has been estimated in this model for improving the performance of overall network. Here, the training complexity and computations of the proposed model have been effectively minimized with the selection of loss function Ls_{SR} as represented in the following model as shown in Eq. (3).

$$Ls_{SR}(F_{\delta}(X_{LRi}), X_{HRi}) = \|F_{\delta}(X_{LRi}) - X_{HRi}\|_1 \qquad (3)$$

The architecture model of the proposed DMResNet model is shown in Fig. 2, which obtains the LR image as the input for processing. Then, it is directly up-scaled to a HR image with the use of DMResNet algorithm, which comprises the main modules of feature extraction and reconstruction.

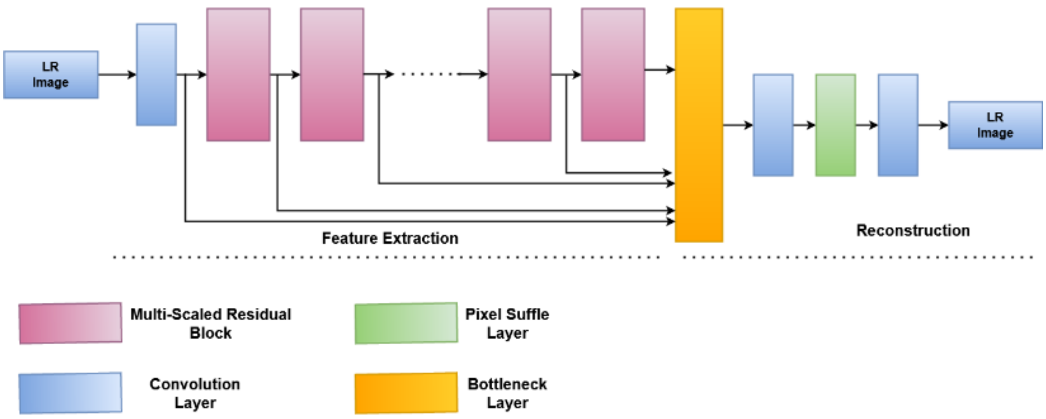


Figure 2. Architecture model of DMResNet.

B. *Multi-scale Feature Extraction*

In the proposed framework, the multi-scale residual block has been added to predict the features of image at varying scales. The architecture model of feature extraction is shown in Fig. 3, which holds the processes of feature fusion and residual learning.

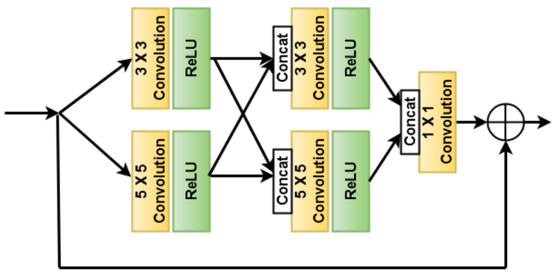


Figure 3. Multi-scaled residual block architecture model.

Here, the bypass network is constructed with the use of various convolutional kernels. Based on this process, the information sharing among the bypass networks is enabled, which helps to efficiently detect the features of the image at varying scales [29, 31]. The bypass operation is performed by using the following models Eq. (4)–(8).

$$T_1 = \rho(\varphi^1_{3\times3} \times Y_{n-1} + \mathfrak{B}_1) \qquad (4)$$

$$R_1 = \rho(\varphi^1_{5\times5} \times Y_{n-1} + \mathfrak{B}_1) \qquad (5)$$

$$T_2 = \rho(\varphi^2_{3\times3} \times [T_1, R_1] + \mathfrak{B}_2) \qquad (6)$$

$$R_2 = \rho(\varphi^2_{5\times5} \times [R_1, T_1] + \mathfrak{B}_2) \qquad (7)$$

$$T' = \varphi^3_{1\times1} \times [T_2, R_2] + \mathfrak{B}_3 \qquad (8)$$

where, $[T_1, R_1]$, $[T_2, R_2]$are the concatenation operations, $\mathfrak{B}$ denotes the bias value, φ represents the weight value, $\rho(a) = \max(0, a)$ is the kernel function of ReLU layer. Consequently, the residual learning is performed to highly improve the efficiency of network, and its multi-scaled residual block is determined by using the following Eq. (9).

$$H_n = T' + H_{n-1} \tag{9}$$

where, H_n is the input, and H_{n-1} represents the output. This type of residual learning could effectively reduce the computational complexity of DMResNet algorithm.

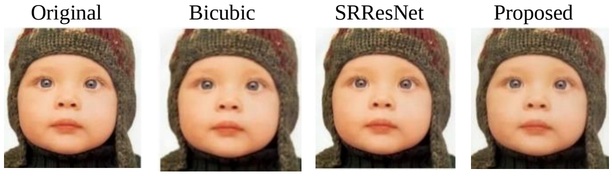
C. *Image Reconstruction*

After feature extraction and residual learning, the operations such as feature fusion and image reconstruction have been performed to generate the HR interpolated image. The features of the input image must be completely utilized and transferred to the network's output for restoration. These traits, however, eventually vanish during transmission as network depth rises. This issue has prompted a number of solutions, the most straightforward and effective of which is the skip connection. These techniques all aim to establish various connections between various layers. However, these techniques can't make full use of the input image's attributes and produce an excessive amount of redundant data that serves no useful purpose. The network's capacity for spatial expression steadily declines while its capacity for semantic expression increasingly rises. Also, each multi-scaled residual block's output has unique characteristics. The quality of the reconstructed images will therefore be strongly impacted by how to utilize these hierarchical characteristics to their best potential. A straightforward hierarchical feature fusion structure is used in this work. The multi-scaled residual block's output is sent to the network's edge for reconstruction. These feature maps have a lot of extraneous features, to say the least. But, if they are used directly for reconstruction, the computational complexity will significantly rise. To solve this problem, the most essential information are extracted from these features with the use of bottleneck layer with the kernel size of 1×1. Then, the output of feature fusion Z_{LR} is modeled as shown in Eq. (10).

$$Z_{LR} = \varphi \times [H_0, H_1, H_2 \dots H_N] + \mathfrak{B} \tag{10}$$

where, $H_0 \dots H_N$ indicates the concatenation operation, and H_0 is the output of 1st convolution layer. Finally, the image reconstruction is performed to generate an interpolated HR image. Fig. 4 shows the original, and SR images generated by the interpolation methods such as bicubic, SRResNet and proposed DMResNet models.

Original | Bicubic | SRResNet | Proposed



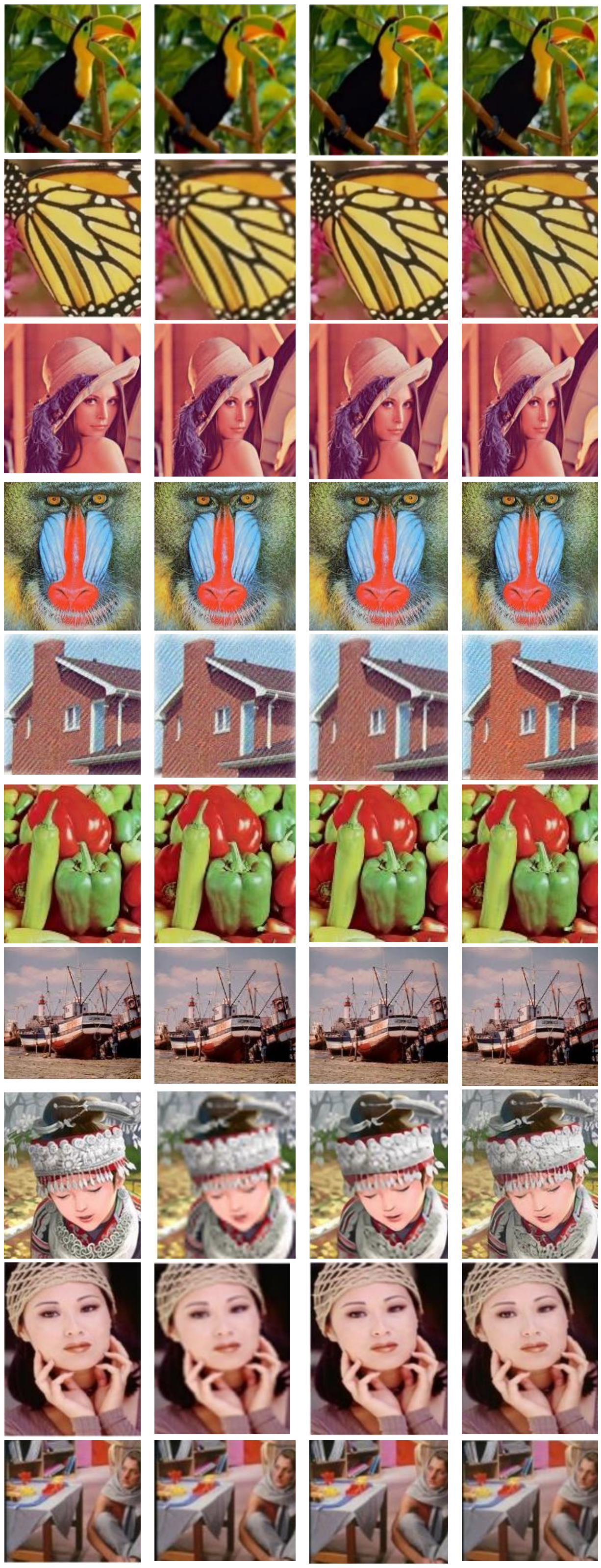


Figure 4. Sample images generated by the interpolation methods.

IV. Result and Discussion

Using a variety of datasets and performance metrics, this section verifies the outcomes and effectiveness of the suggested DMResNet-based picture interpolation process. This analysis uses the popular benchmarking datasets such as IAPR TC-12, DIV 2K and CVDS. In order to validate the results, the parameters such as Peak Signal to Noise Ratio (PSNR), run time, Structural Similarity Index Measure (SSIM), Estimated Peak Signal to Noise Ratio (EPSNR) and Feature Similarity Index Measure (FSIM) have been considered. The last row of the Tables II–IV used in this analysis with the title of average can be utilized to exhaustively make a comparison between the proposed technique and the existing Bi-Cubic & AFFSR approaches based on the parameters of run time, PSNR, and SSIM.

A. Run Time Analysis of Bench Mark Images

The run time(s) analysis of benchmark images are compared between Bi-Ciubic, AFFSR (Adaptive Fractional Fourier Super resolution) and proposed (DMResNet) methods are shown as tabulation in Table II and Graphical representation in Fig. 5. The comparison findings shown in Table II and Fig. 5 demonstrates the proposed DMResNet provides an improved performance over the conventional Bi-Cubic and AFFSR approaches [32] with reduced run time. The average run time of the proposed technique is significantly lower than the Bi-Cubic and AFFSR methods. This improved value shows that the proposed method outperforms the Bi-Cubic method in terms of image quality. The average run time of the proposed technique is significantly lower than the Bi-Cubic and AFFSR methods. A well-presented results section coupled with a convincing discussion will definitely prove the novelty and importance of your study. It should provide a concise and precise description of the experimental results, their interpretation, as well as the experimental conclusions that can be drawn.

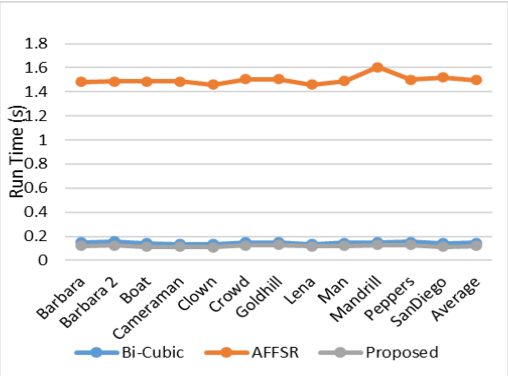


Figure 5. Comparative analysis based on run time.

TABLE II. RUN TIME (S) ANALYSIS

Images	Bi-Cubic	AFFSR	Proposed
Barbara	0.1519	1.4812	**0.1214**
Barbara 2	0.1563	1.4864	**0.1254**
Boat	0.1420	1.4855	**0.1136**
Cameraman	0.1364	1.4849	**0.1124**
Clown	0.1357	1.4597	**0.1109**
Crowd	0.1499	1.5036	**0.1233**
Goldhill	0.1518	1.5063	**0.1267**
Lena	0.1352	1.4590	**0.1147**
Man	0.1473	1.4886	**0.1195**
Mandrill	0.1518	1.6052	**0.1265**
Peppers	0.1521	1.5011	**0.1274**
SanDiego	0.1442	1.5177	**0.1139**
Average	0.1462	1.4983	**0.1196**

B. Comparison of PSNR and SSIM of Benchmark Images

The PSNR and SSIM values of benchmark images are validated and compared between Bi-Ciubic, AFFSR and proposed (DMResNet) methods in Table III–Table IV and its graphical representation is shown in Figs. 6–7. The PSNR of the existing and proposed methods are validated and compared using different images. The analysis indicate that the proposed method performs better than the previous methods in the majority of the images.

TABLE III. PSNR ANALYSIS

Images	Bi-Cubic	AFFSR	Proposed
Barbara	25.3511	25.567	**28.146**
Barbara 2	27.7831	28.7662	**30.457**
Boat	29.9472	30.9054	**32.987**
Cameraman	35.7448	38.3443	**40.234**
Clown	32.6602	34.1681	**36.478**
Crowd	32.6606	34.2721	**36.364**
Goldhill	31.4538	32.2247	**34.229**
Lena	34.1194	35.4265	**37.114**
Man	31.053	32.04	**34.987**
Mandrill	23.6416	24.3146	**26.589**
Peppers	31.855	32.6989	**34.897**
SanDiego	30.1996	31.3127	**33.478**
Average	30.5394	31.6700	**33.83**

TABLE IV. SSIM ANALYSIS

Images	Bi-Cubic	AFFSR	Proposed
Barbara	0.6656	0.7278	**0.7465**
Barbara 2	0.7703	0.8228	**0.8498**
Boat	0.6639	0.7150	**0.7341**
Cameraman	0.8410	0.8874	**0.9068**
Clown	0.7368	0.7740	**0.7954**
Crowd	0.8749	0.9027	**0.9247**
Goldhill	0.7281	0.7778	**0.7936**
Lena	0.7174	0.7633	**0.7844**
Man	0.7632	0.8108	**0.8347**
Mandrill	0.6393	0.7242	**0.7468**
Peppers	0.6527	0.6968	**0.7198**
SanDiego	0.6633	0.7447	**0.7635**
Average	0.7264	0.7789	**0.8008**

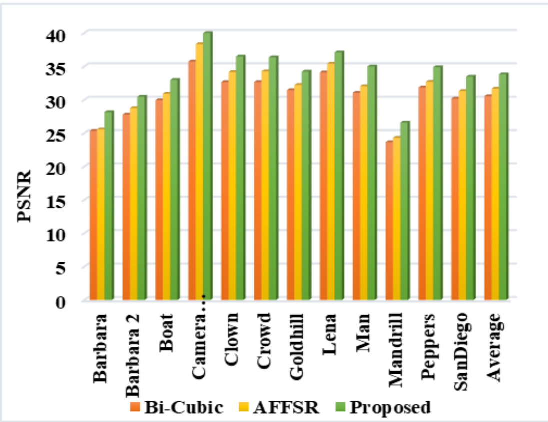


Figure 6. Comparative analysis based on PSNR.

In terms of the quality of the created image, the average row's conclusion shows that the proposed approach outperforms the other models with improved PSNR value. According to the analysis, it is concluded that the proposed DMResNet algorithm overwhelms the existing Bi-cubic and AFFSR techniques with improved run time, PSNR and SSIM results.

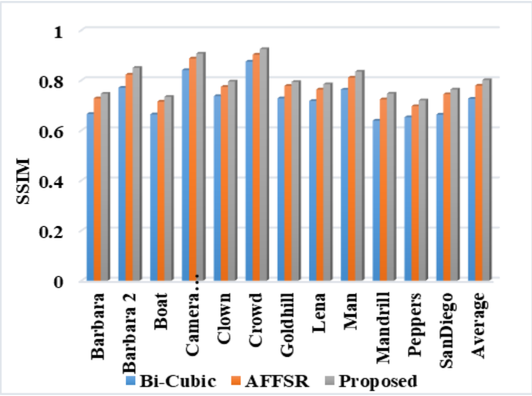


Figure 7. Comparative analysis based on SSIM.

C. *Comparison Results of RSME and FSIM While Using Dataset Images*

The RMSE and FSIM analysis of dataset images are compared between Nearest Neighbor (NN), Bilinear (BL), Bi-Cubic Spline (BC), Bi-Cubic Hermite (BH), Karim and Saaban (KS), Bernstein-Bezier (BB), and proposed (DMResNet) methods are shown as tabulation in Table V and Table VI and Graphical representation in Figs. 8 and 9. Table V and Fig. 8 validates the Root Mean Square Error (RMSE) of the conventional [30] and proposed interpolation mechanisms. Moreover, we discovered that the proposed technique provided reduced values for the majority of images in terms of RMSE. According to the obtained results through simulation, it is discovered that the optimal parametric values that would yield the greatest outcomes. In addition, Table VI presents the comparative analysis among the existing and proposed interpolation methods based on the parameter of FSIM.

TABLE V. RMSE ANALYSIS

Images	**NN**	**BL**	**BC**	**BH**	**KS**	**BB**	**Proposed**
Fishing boat	2.71	2.84	2.68	2.64	2.64	2.59	2.14
Baboon	3.22	3.30	3.11	3.14	3.14	3.07	2.48
Couple	2.93	3.03	2.87	2.86	2.86	2.81	2.09
Watch	2.64	2.75	2.60	2.59	2.59	2.55	2.13
Man	2.93	3.03	2.83	2.81	2.81	2.76	2.17
Girl	2.12	2.20	2.08	2.81	2.81	2.76	2.08
City	3.51	3.62	3.45	3.46	3.46	3.42	2.31
Sails	3.37	3.53	3.34	3.34	3.34	3.28	2.35
Dome	2.29	2.39	2.28	2.23	2.23	2.20	1.45
House	2.56	2.68	2.54	2.52	2.52	2.48	1.36
Average	2.82	2.93	2.77	2.84	2.84	2.79	2.05

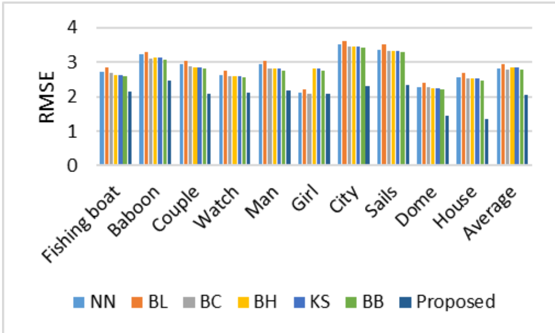


Figure 8. Comparative analysis based on RMSE.

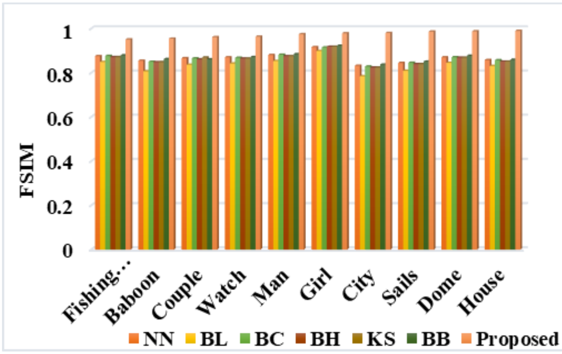


Figure 9. Comparative analysis based on FSIM.

TABLE VI. FSIM ANALYSIS

Images	**NN**	**BL**	**BC**	**BH**	**KS**	**BB**	**Proposed**
Fishing boat	0.8754	0.8489	0.8763	0.8713	0.8713	0.8790	0.9509
Baboon	0.8546	0.8063	0.8503	0.8483	0.8486	0.8622	0.9547
Couple	0.8660	0.8367	0.8653	0.8610	0.8697	0.8610	0.9614
Watch	0.8704	0.8421	0.8679	0.8649	0.8649	0.8715	0.9635
Man	0.8808	0.8535	0.8818	0.8755	0.8755	0.8841	0.9748
Girl	0.9162	0.8989	0.9149	0.9184	0.9184	0.9227	0.9787
City	0.8321	0.7848	0.8284	0.8238	0.8238	0.8372	0.9804
Sails	0.8448	0.8091	0.8452	0.8396	0.8396	0.8509	0.9865
Dome	0.8700	0.8449	0.8705	0.8685	0.8685	0.8770	0.9878
House	0.8576	0.8312	0.8569	0.8508	0.8508	0.8588	0.9899
Average	0.8668	0.8356	0.8658	0.8622	0.8631	0.8704	0.9729

Based on the analysis, we discovered that the proposed scheme is superior to the other schemes for all images. The Proposed model DMResNet outperforms the state-of-art interpolation techniques in its performance, by yielding an increase of 8% in PSNR, 6% in SSIM, 1.2% in FSIM, and a decrease of 38.79% in RMSE, 5.875 times in run time analysis.

Table VII shows the ablation study of the proposed model, where the training and testing operations are performed based on the luminance channel with the upscaling factors of ×2, ×3, ×4 and ×8.

TABLE VII. ABLATION STUDY

Layers	Input	Kernel Size	Output
Convolution_input	64	3×3	64×2×2
Pixel_shuffle (×2)	64×2×2	-	64
Convolution_output	64	3×3	1
Convolution_input	64	3×3	64×3×3
Pixel_shuffle (×3)	64×3×3	-	64
Convolution_output	64	3×3	1
Convolution_input	64	3×3	64×4×4
Pixel_shuffle (×4)	64×4×4	-	64
Convolution_output	64	3×3	1
Convolution_input	64	3×3	64×8×8
Pixel_shuffle (×8)	64×8×8	-	64
Convolution_output	64	3×3	1
Convolution_input	64	3×3	64×M×M
Pixel_shuffle (×M)	64×M×M	-	64
Convolution_output	64	3×3	1

TABLE VIII. QUANTITATIVE ANALYSIS

Algorithms	Scale	PSNR	SSIM
Bicubic	(×2)	33.69	0.9284
CNN	(×2)	37	0.9560
Super Resolution—CNN	(×2)	36.71	0.9536
Proposed	**(×2)**	**38.15**	**0.9632**
Bicubic	(×3)	30.41	0.8654
CNN	(×3)	33.21	0.9148
Super Resolution—CNN	(×3)	32.47	0.9024
Proposed	**(×3)**	**34.70**	**0.9298**
Bicubic	(×4)	28.45	0.8032
CNN	(×4)	30.68	0.8647
Super Resolution—CNN	(×4)	30.51	0.8536
Proposed	**(×4)**	**32.49**	**0.9032**
Bicubic	(×8)	24.41	0.6153
CNN	(×8)	25.74	0.6748
Super Resolution—CNN	(×8)	25.36	0.6472
Proposed	**(×8)**	**26.58**	**0.8894**

Moreover, Table VIII presents the quantitative comparative analysis of the existing Bi-cubic, CNN, Super resolution CNN models and proposed DMResNet models based on the parameters of PSNR and SSIM at varying scales such as ×2, ×3, ×4 and ×8. The findings indicate that the proposed model outperforms the other interpolation techniques with improved performance for all scaling factors.

V. CONCLUSION

This paper presents a new image interpolation framework with the use of advanced deep learning algorithm, called as, DMResNet. The contribution of this paper is to utilize a novel deep learning model for improving image super resolution. The main purpose of this paper is to generate an interpolated HR image from the LR image with the use of deep learning algorithm. Adaptively detecting the image features at various scales is accomplished in this research by the use of a multi-scale residual block that is reliable and effective. We presented a multi-scale residual network based on this concept, where the localized multi-scale features and the recursive features are used to construct an effective SR model, resulting in an accurate SR image. Moreover, we obtained an encouraging results by using the multi-scaled residual blocks module for other computer vision tasks like image de-noising and de-hazing. For validation, several popular benchmarking datasets are utilized in this work, where the results are assessed using different parameters such as run time, PSNR, RMSE, SSIM, and FSIM. The results indicate that the proposed method outperforms the cutting-edge techniques and produces an effective results by using several datasets.In future, the present work can be enhanced by using an advanced interpolation model for the reconstruction of underwater images.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Diana Earshia V has been involved in organizing the paper, contributing to data analysis, making English corrections and grammar checks, as well as assisting in deriving the mathematical equation. On the other hand, Sumathi M has contributed to the background study of the paper, assisted in mathematical derivations, provided a factual review, and helped edit the manuscript. Their combined efforts have likely played a crucial role in the development and improvement of the paper; all authors had approved the final version.

ACKNOWLEDGMENT

I am grateful to all of those with whom I have had the pleasure to work during this and other related research work.

REFERENCES

[1] S. Kim, D. Jun, B. G. Kim, H. Lee, and E. Rhee, "Single image super-resolution method using CNN-based lightweight neural networks," *Applied Sciences*, vol. 11, 1092, 2021.

[2] K. W. Hung, K. Wang, and J. Jiang, "Image interpolation using convolutional neural networks with deep recursive residual learning," *Multimedia Tools and Applications*, vol. 78, pp. 22813–22831, 2019.

[3] J. Zheng, W. Song, Y. Wu, and F. Liu, "Weighted direct nonlinear regression for effective image interpolation," *IEEE Access*, vol. 7, pp. 8646–8659, 2019.

[4] K. Zhang, H. Hu, K. Philbrick, G. M. Conte, J. D. Sobek, P. Rouzrokh *et al.*, "SOUP-GAN: Super-resolution MRI using generative adversarial networks," *Tomography*, vol. 8, pp. 905–919, 2022.

[5] J. Liu, Y. Liu, H. Wu, J. Wang, X. Li, and C. Zhang, "Single image super-resolution using feature adaptive learning and global structure sparsity," *Signal Processing*, vol. 188, 108184, 2021.

[6] M. Jalali, H. Behnam, F. Davoodi, and M. Shojaeifard, "Temporal super-resolution of 2D/3D echocardiography using cubic B-spline interpolation," *Biomedical Signal Processing and Control*, vol. 58, 101868, 2020.

[7] Z. Xiuyun, L. Huacai, and Z. Chenyan, "Super resolution image restoration algorithm: Based on wavelet and interpolation," in *Proc. 2020 3rd World Conference on Mechanical Engineering and Intelligent Manufacturing (WCMEIM)*, pp. 95–98, 2020.

[8] J. Zhang, M. Shao, L. Yu, and Y. Li, "Image super-resolution reconstruction based on sparse representation and deep learning," *Signal Processing: Image Communication*, vol. 87, 115925, 2020.

[9] Y. Chen, L. Liu, V. Phonevilay, K. Gu, R. Xia, J. Xie *et al.*, "Image super-resolution reconstruction based on feature map attention mechanism," *Applied Intelligence*, vol. 51, pp. 4367–4380, 2021.

[10] J. Yang, S. Shen, H. Yue, and K. Li, "Implicit transformer network for screen content image continuous super-resolution," *Advances in Neural Information Processing Systems*, vol. 34, pp. 13304–13315, 2021.

[11] S. S. Rajput and K. Arya, "A robust face super-resolution algorithm and its application in low-resolution face recognition system," *Multimedia Tools and Applications*, vol. 79, pp. 23909–23934, 2020.

[12] X. Yao, Q. Wu, P. Zhang, and F. Bao, "Weighted adaptive image super-resolution scheme based on local fractal feature and image roughness," *IEEE Transactions on Multimedia*, vol. 23, pp. 1426–1441, 2020.

[13] D. Qiu, S. Zhang, Y. Liu, J. Zhu, and L. Zheng, "Super-resolution reconstruction of knee magnetic resonance imaging based on deep learning," *Computer Methods and Programs in Biomedicine*, vol. 187, 105059, 2020.

[14] M. Rostami, U. Muhammad, S. Forouzandeh, K. Berahmand, V. Farrahi, and M. Oussalah, "An effective explainable food recommendation using deep image clustering and community detection," *Intelligent Systems with Applications*, vol. 16, 200157, 2022.

[15] Y. Zou, L. Zhang, C. Liu, B. Wang, Y. Hu, and Q. Chen, "Super-resolution reconstruction of infrared images based on a convolutional neural network with skip connections," *Optics and Lasers in Engineering*, vol. 146, 106717, 2021.

[16] F. Ling and G. M. Foody, "Super-resolution land cover mapping by deep learning," *Remote Sensing Letters*, vol. 10, pp. 598–606, 2019.

[17] S. M. A. Bashir, Y. Wang, M. Khan, and Y. Niu, "A comprehensive review of deep learning-based single image super-resolution," *Peer J. Computer Science*, vol. 7, p. e621, 2021.

[18] J. Jakob, M. Gross, and T. Günther, "A fluid flow data set for machine learning and its application to neural flow map interpolation," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, pp. 1279–1289, 2020.

[19] G. Pandey and U. Ghanekar, "A hybrid single image super-resolution technique using fractal Interpolation and convolutional neural network," *Pattern Recognition and Image Analysis*, vol. 31, pp. 18–23, 2021.

[20] X. Huo, R. Tang, L. Ma, K. Shao, and Y. Yang, "A novel neural network for super-resolution remote sensing image reconstruction," *International Journal of Remote Sensing*, vol. 40, pp. 2375–2385, 2019.

[21] J. Hou, Y. Si, and X. Yu, "A novel and effective image super-resolution reconstruction technique via fast global and local residual learning model," *Applied Sciences*, vol. 10, 1856, 2020.

[22] P. Garima and G. Umesh, "A hybrid single image super-resolution technique using fractal interpolation and convolutional neural network," *Pattern Recognition and Image Analysis (Advances in Mathematical Theory and Applications)*, vol. 31, pp. 18–23, 2021.

[23] M. Ismail, J. Yang, C. Shang, and Q. Shen, "Image super resolution with sparse data using ANFIS interpolation," in *Proc. 2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 2020, pp. 1–7.

[24] Y. Zhou, Y. Zhang, X. Xie, and S. Y. Kung, "Image super-resolution based on dense convolutional auto-encoder blocks," *Neurocomputing*, vol. 423, pp. 98–109, 2021.

[25] W. Wang, J. Hu, X. Liu, J. Zhao, and J. Chen, "Single image super resolution based on multi-scale structure and non-local smoothing," *EURASIP Journal on Image and Video Processing*, vol. 2021, 16, 2021.

[26] M. Ismail, J. Yang, C. Shang, and Q. Shen, "Single frame image super resolution using ANFIS interpolation: An initial experiment-based approach," in *Proc. Advances in Computational Intelligence Systems: Contributions Presented at the 19th UK Workshop on Computational Intelligence, September 4–6, 2019, Portsmouth*, UK, 2020, pp. 27–40.

[27] Z. Hui, X. Gao, Y. Yang, and X. Wang, "Lightweight image super-resolution with information multi-distillation network," in *Proc. the 27th ACM International Conference on Multimedia*, 2019, pp. 2024–2032.

[28] O. Aydin and R. G. Cinbiş, "Single-image super-resolution analysis in DCT spectral domain," *Balkan Journal of Electrical and Computer Engineering*, vol. 8, pp. 209–217, 2020.

[29] J. Ji, B. Zhong and K. K. Ma, "Image interpolation using multi-scale attention-aware inception network," *IEEE Transactions on Image Processing*, vol. 29, pp. 9413–9428, 2020.

[30] C. H. Yeh, C. H. Lin, M. H. Lin, L. W. Kang, C. H. Huang, and M. J. Chen, "Deep learning-based compressed image artifacts reduction based on multi-scale image fusion," *Information Fusion*, vol. 67, pp. 195–207, 2021.

[31] X. Feng, X. Li, and J. Li, "Multi-scale fractal residual network for image super-resolution," *Applied Intelligence*, vol. 51, pp. 1845–1856, 2021.

[32] A. Faramarzi, A. Ahmadyfard, and H. Khosravi, "Adaptive image super-resolution algorithm based on fractional fourier transform," *Image Analysis and Stereology*, vol. 41, pp. 133–144, 2022.