SENExt

by Wazir Muhammad

Submission date: 26-Sep-2022 10:10PM (UTC+0500)
Submission ID: 1909575683
File name: SENext_Final_Version.docx (3.95M)
Word count: 8455
Character count: 46698

SENext: Squeeze-and-ExcitationNext for Single Image Super-Resolution

Wazir Muhammad[a], Supavadee Aramvith[b,*] and Takao Onoye[c]

[a] *Department of Electrical Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, 10330, Thailand*
[b*] *Multimedia Data Analytics and Processing Unit, Department of Electrical Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok 10330, Thailand.*
[c] *Graduate School of Information Science and Technology, Osaka University, 1-5 Yamada-Oka, Suita, 565-0871 Japan.*

ABSTRACT

Recent research on single image super-resolution (SISR) using deep convolutional neural networks (CNNs) has shown significant development in the area computer vision-based tasks specially image and video processing. SISR seeks to reconstruct a visibly appealing high-quality / high-resolution (HR) output image from a low-quality / low-resolution (LR) input image as its primary goal. However, most existing CNN-based image super-resolution (SR) frameworks often use a deeper and broader network architecture that requires a sizeable computational resource, risk of overfitting, increases computational complexity, and more memory consumption, as well as takes more processing time during the evaluations. To resolve these problems, we propose a Squeeze-and-ExcitationNext for Single Image Super-Resolution concept named as SENext. In detail, the squeeze-and-excitation blocks (SEB) are used in our network architecture to reduce the computational cost and adopt the channel-wise feature mappings to adaptively recalibrate the features. Furthermore, local, sub-local and global skip connections are employed between each SEB to enable the feature reusability and stabilize training convergence smoothly. Instead of hand-designed bicubic upsampling at pre-processing step, we perform post-upsampling at the later end to reconstruct the high-resolution (HR) image. Extensive quantitative and qualitative experiments are performed on the benchmark test dataset, including Set5, Set14, BSDS100, Urban100, and Manga109. These experimental evaluations validate the superiority of the SENext over other deep CNN image SR methods in terms of PSNR/SSIM, FLOPs, Number of parameters, processing speed, and visually pleasing effect.

Keywords: Convolutional Neural Networks; LeakyReLU activation Function, Squeeze-and-excitation block.

---

1. Introduction

One of the most significant research areas in deep learning and image processing is a single image super-resolution (SISR). Reconstructing the visually appealing high-resolution (HR) output image from the low-resolution (LR) input image is the primary function of SISR. However, SISR is still a difficult task and is considered an inverse ill-posed problem because numerous algorithms [1-5] have been suggested. Still, performance is not satisfactory and has more computational complexity. Recently, deep convolutional neural networks (CNNs) captured the market for image SR, and the research community shifted from the old hand-designed approach to a newly deep CNN-based approach. Initially, Dong et al. proposed a shallow type Super-Resolution Convolutional Neural Network (SRCNN) [6] architecture to

*Corresponding author. Tel.: +6622186911.
E-mail address: supavadee.a@chula.ac.th.

reconstruct a better HR image from the bicubic interpolated generated LR input image [6]. Compared with the earlier conventional approaches, SRCNN [6] can improve performance through its shallow network architecture when reconstructing the HR image. SRCNN [6] consists of three basic types of CNN layers: patch extraction, mapping, and reconstructed layers. Apart from the success of SRCNN [6] in the image super-resolution, it has many shortcomings, including slow training speed, poor real-time reconstruction, bicubic interpolation stage as a pre-processing stage, and large convolution kernels used during the model design. In response to these problems, the same author has proposed the revised version of SRCNN [6] and replaced the bicubic interpolation with a learnable upsampling (transpose convolution) layer to accomplish post-upsampling SR named Fast Super-Resolution CNN (FSRCNN) [7]. Furthermore, larger kernel sizes of SRCNN [6] are replaced with small convolution kernels to optimize the efficiency of training and reconstruction. FSRCNN [7] improved the performance and decreased the computational cost compared to the previous SRCNN [6]. The main drawback of FSRCNN [7] is the capacity of a network is limited. Following the concept of the Visual Geometry Group network (VGG-net) [8] that was used for ImageNet classification, Kim et al. first time introduced the idea of very deep super-resolution (VDSR) [9], which pushed up the network and serially stacking multiple layers up to 20 layers. The performance of the VDSR [9] model significantly improved over previous models. This method suggested that deeper model architecture is the better architecture to increase the visual quality of the HR image. Initially, sub-pixel layer-based model used in image super-resolution suggested by Shi et al. and named as an Efficient Sub-pixel Convolutional Neural Network (ESPCN) [10] to decrease the computational burden as well as revise the upscaling process. In this approach, the authors change the pre-stage upscaling bicubic operator with a sub-pixel convolution layer, and features are recovered from the original low-dimensional space to decrease the model processing time of during the training as well as testing. Kim et al. suggested the new way of architecture known as Deeply Recursive Convolutional Network for image super-resolution (DRCN) [11] and replaced the serial way of a combination of CNN layers with a recursive manner. This architecture's main benefit is to constantly maintain network parameters, but the training convergence process is too slow. Additionally, to obtain better reconstruction performance, the SR models used the concept of a deeper model and stacking the side layer by the side. In some cases, the model depth increases up to 100 layers observed [12]. A super-resolution model's performance can be enhanced by increasing its spatial depth, but doing so will suffer a significant computational expense and memory usage. To lessen the computational complexity and increase the processing speed of image SR models inspired by the SENet [13] and SESR [14], we proposed a Squeeze-and-ExcitationNext for a single image super-resolution named SENext. In our SENext method, squeeze-and-excitation block (SEB) is used to develop the interdependencies between respective channels and reweight the new features. Additionally, as shown in Fig. 1, a bicubic pre-processing operation is employed as an upscaling factor to rebuild the HR image using state-of-the-art methods such as SRCNN [6], VDSR [9], and DRCN [11]. The main issue with these approaches having a more computational cost and reconstructing HR images is introducing blurry results. We replaced the initial feature extraction layers with a feature extraction block (FEB) to resolve these issues. A single-stage block was replaced with a two-stage squeeze and excitation block (SEB) to reconstruct the visually pleasing HR image with a low computational cost.

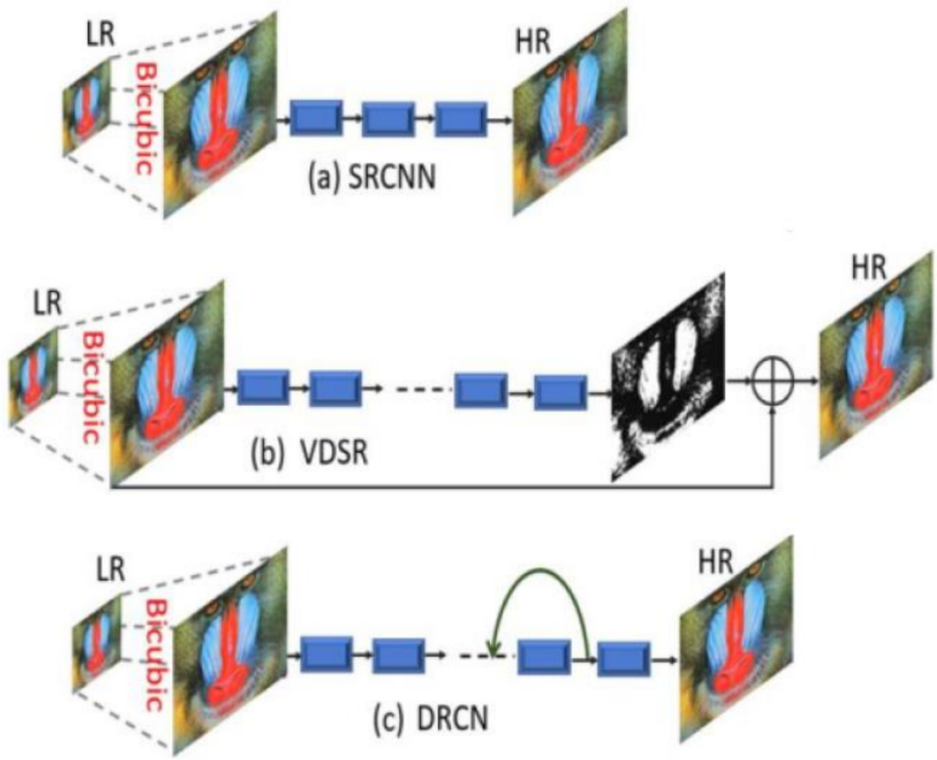


Fig. 1. Pre-processing interpolation-based image super-resolution architectures of SRCNN [6], VDSR [9], and DRCN [11].

Furthermore, single local skip connection-based image super-resolution approaches face the loss of feature information at the later end of the layers and work as a dead layer. This issue introduces the vanishing gradient problem occurring in training [8, 15, 16], our proposed model handle this issue with the support of global as well as local skip connections. In addition, selecting the proper activation function is crucial for developing deep CNN methods. Rectified Linear Units (ReLU) are currently the most popular activation function. As Krizhevsky et al. [8], the advantages of using the ReLU activation function include faster training speed and decreased saturation problems. Still, several recent papers address the issues of exploding (i.e., retraining too much information) or dying (i.e., retaining too little information) during the training [8, 15, 16]. It is desirable to suggest a novel activation function to address the abovementioned shortcomings. In contrast to ReLU and PReLU activation functions, the novel nonlinear activation function proposed in this work is a LeakyReLU.

The main contribution of our proposed method is as under:

- To reduce the computational cost and obtain faster convergence during the training phase, we replace standard ResNet blocks with squeeze and excitation (SEB) blocks inspired by the Squeeze and Excitation networks. Compared to other image SR methods, our suggested model outperforms them by a factor of 2×, 3×, 4×, and 8× benchmark not only in terms of speed but also in terms of computational cost.

- The deeper model faces the problems of Dying Rectified Linear Unit (ReLU), which means the condition in which many ReLU neurons send output values as zero, and the whole network gets stuck and never improves the performance. We replace the ReLU with the LeakyReLU to initiate the dead features introduced by zero gradients.
- The single local and global skip connection does not reconstruct the visually pleasing high-quality HR image and introduces blurry artifacts to the HR output image. We adopt an different approach and extracted the features information from the multi-local, sub-local, and global skip connections to reconstruct the visually pleasing, high-quality HR image.

The remaining sub-section of our work is explained under. Section 2 discusses the related works of deep CNN image SR methods. Section 3 explains the designed framework for SISR in detail. In section 4, we discussed the experimental evaluations with other state-of-the-art methods. Finally, section 5 explains the conclusion part.

2. Related works

The objective of SISR is to reconstruct the original LR input image into a visually appealing HR output image that contains detailed information. Many researchers have started solving the image SR problem differently since deep CNN learning-based architecture became famous. In this paper, we only go into detail about current deep learning CNN-based approaches. The Super-Resolution Convolutional Neural Network (SRCNN) [6] is the first deep learning-based solution to the SISR problem proposed by Dong et al. Comparing this strategy to all earlier SR techniques, and it exhibits considerable gains. SRCNN [6] model depends on three CNN layers to predict the output from the interpolated version of the upscaled image to reconstruct the HR image. Although, there is some weakness in this model. First, the proposed model used bicubic interpolation to upscale the original LR image, but bicubic interpolation introduced blurry results and did not design for this purpose. Second, image reconstruction information is still not satisfactory. The third is the slow convergence rate which takes more training time. Wang et al. [17] introduce the sparse prior network for reconstructing the HR image, known as the Sparse Coding Network (SCN) [17]. The computational performance of SCN is also improved than earlier SR methods from SRCNN [6] as well. Wang et al. further modified the model and replaced the non-linear mapping with a set of coding sparse sub-networks [18]. The main disadvantage of SCN [17] network architecture is the higher computational cost, leading to many problems in real-time applications.

To speed up the reconstruction process of image super-resolution, Dong et al. introduced the Fast Super-Resolution Convolutional Neural Network (FSRCNN) [7] architecture. FSRCNN [7] is an upgraded and faster version of the SRCNN [6] design. Its straightforward network design uses one deconvolution layer and four CNN layers to upsample the original input LR images without using interpolation techniques. Compared to SRCNN [7] performs better and has lower computational complexity. Still, it has a smaller network capacity. Efficient sub-pixel convolutional neural network (ESPCN) [10] is a simple, efficient, and fast image super-resolution method, that can apply on real-time image and video applications.

A very deep SR (VDSR) network with residual skip connection was introduced by Kim et al. [9], which was modeled after the Visual Geometry Group network (VGG-net) used in the ImageNet for classification [8]. Utilizing the 20 CNN trainable layers, the VDSR [9] network exhibited considerable performance and improvement over the SRCNN [6] and FSRCNN [7] networks. The global residual learning connection is used with the support of a faster convergence rate to lower the training complexity of VDSR. Though, VDSR [9] method uses the bicubic interpolation-based upscaled type of input image rather than the actual pixel values, which results in increased memory usage and high computational costs. In addition, Kim et al. presented a Deeply Recursive Convolutional Network (DRCN) [11] for image SR framework that employs several convolution layers. The key advantage of DRCN [11] is that it has constant training parameters (number of parameters). Although there are more recursions, the main drawback of DRCN [11] is that it slows the process of training convergence. The authors also applied the skip connection recursively to enhance model performance. The Residual Encoder-Decoder Networks (RED) are a notion that Mao et al. extend and proposed the RED [19] model and uses residual learning with symmetric convolution operation, obtaining better performance. As a result, these findings support the idea that "the Deeper the Better." Contrarily, a shallow and deeper, fast deep learning-based approach was proposed by Romano et al., named Rapid and Accurate Image Super-Resolution (RAISR) [20]. In this approach, the author classifies the input image patches concerning the angle of patches, coherence, and strength to learn the mappings from the original LR image to reconstruct the HR image. To rebuild the HR image, Lai et al. developed a deep Laplacian Pyramid Super-Resolution Network (LapSRN) [21], a novel image SR design. The LapSRN [21] architecture is based on many pyramid layers, each of which has a deconvolution layer acting as an upsample. Denoising convolutional neural networks (DnCNNs) were suggested by Zhang et al. [22] to speed up the development of an extremely deep convolutional neural network design. The DnCNN network stacks convolutional neural networks with batch normalisation (BN) layers prior to the ReLU activation function, just like the SRCNN [6] network. Despite producing positive results, the model is computationally expensive because it uses a batch normalization layer. A progressive upsampling network is the more adaptable scaling factor suggested by Zhao et al. [23] named a gradual upsampling network (GUN). GUN performs forward and backward computations during the training to upscale the features. The 52 CNN layers with recursive residual networks were first time suggested by Tai et al. [24]. Ledig et al. [25] use a deep CNN with residual skip connections having 16 blocks to recover the upsampled version of output image. Lim et al. [26] suggested an improved deep super-resolution network architecture to boost the model's training effectiveness and win the NTIRE2017 SR challenge [27]. Tai et al. proposed the deepest model for image restoration, a persistent memory network (MemNet), which layers a number of memory blocks to create persistent memory [28]. MemNet consists of cascaded memory blocks, which fuse the global features.

Yamanaka et al. [29] developed a deep convolutional neural network-based framework for image SR and suggested combining parallelized CNN layers and skip connections. The two networks they use most frequently are SR image reconstruction network and a feature extraction network for extracting features from various levels. Compared to VDSR [9], this model is shallower. Han et al. proposed a Dual-State Recurrent Network (DSRN), which transmits information from the LR image state to HR image state [30]. They update the signal information at each step before forwarding it to the HR state. A multi-scale residual network (MSRN) was created by Li et al. [31] developed a multi-scale residual network (MSRN), which acquiring the features fusion at various sizes by employing an adaptive feature detection strategy. This method utilized the full hierarchical-based feature type information to recreate the super-resolved HR image. Ahn et al. [32] methods for handling multi-scale information and learning residuals in LR feature space to select appropriate routes [32]. Furthermore, this method provides modules for scale-specific upsampling type with multiple shortcut connections. Choi et al. [33] used the idea of a recursive neural network and proposed a fast and efficient image

SR with Block State-based Recursive Network (BSRN). This type of network architecture tracks current information status for image features. Zhang et al.[34] proposed the super-resolution network for multiple degradations (SRMD), which reconstructs the HR image by concatenating a LR image with its degradation mapping type. Furthermore, SRMD also designed another fine-tuning-based architecture. Noise-free degraded version of SRMD is named SRMDNF [34]. Multi-scale inception-based super-resolution (MSISRD) method was proposed by Muhammad et al. [35] and before utilizing the inception block to reconstruct the multi-scale feature information for image SR, the authors of this method employ the concept of asymmetric convolution operation to reduce the model's computational cost. Wang et al. [36] demonstrated a dilated convolution neural network that was designed to expand the receptive field without expanding the kernel. Under this approach, a shallow network architecture only increased the size of the receptive field. Twelve layers are used in the Dilated Convolutional Network for SR (DCNSR) to efficiently extract contextual data. The End-to-End Image SR architecture provided short and long-range multi-scale information and substituted a transposed CNN layer for bicubic interpolation upsampling in the HR image reconstruction process [37]. Yang et al. [38] proposed a transposed layer-based network architecture with large-scale components known as a deep recurrent fusion network (DRFN). Su et al. [39] suggested a unique type structure, which entails several sub-networks to gradually reconstruct the HR image. The LR feature map will be utilized as the input for each sub-network, and the output of the transposed convolution will be combined with the residuals to produce the finer one. In image super-resolution, arbitrary enlargement factor is a challenge in real-time applications. Hui et al. [40] introduced an information multi-distillation network (IMDN) that was lightweight. Cascaded information multi-distillation blocks (IMDB), which include components for selective fusion and distillation, were utilized in IMDN. Using an information distillation network (IDN), IMDN also solves the problem of memory consumption and computational cost. Lim et al. [26] produced cutting-edge results by utilizing the residual blocks to construct an extremely broad and deep network architecture known as an enhanced deep super-resolution network (EDSR). Both EDSR and EDSR-baseline were released by the author of EDSR. Hung et al. [41] proposed the architecture of a super-sampling network (SSNet) and used image SR with depthwise separable convolution. This architecture use of the depthwise separable convolution method, which reduces the number of parameters and multiplication operations significantly. Barzegar et al. [42] suggested the modest framework to avoid the training's issue in the deeper network architecture. Multi-scale Xception-based depthwise separable convolution for single image super-resolution (MXDSIR) was proposed by Muhammad et al. [43]. The authors employed a depth-wise separable convolution technique in this paper to reduce computational complexity. Hsu et al. [44] were motivated by a capsule neural network to extract additional possible feature information for image SR. In this study, the authors created two networks for image SR: the Capsule Attention and Reconstruction Neural Network (CARNN) and the Capsule Image Restoration Neural Network. For SR objectives and to learn the features information at various phases, Liu et al. [45] presented a new hierarchical convolutional neural network (HCNN) architecture. The HCNN method involves a three-step hierarchical procedure based on the edge branch extraction, the edge reinforcement branch, and the SR image reconstruction branch. Yang et al. proposed a non-linear perceptual multi-scale network architecture abbreviated as NLPMSNet [46]. In this approach, the author fuses the information of multi-scale image information in a non-linear manner and also uses a cascading-based multi-scale global mechanism to capture the non-local feature information. Reduce the computational cost and as well as more memory consumption. Xiao et al. [47] introduced the idea of powerful lightweight multi-scale feature extraction super-resolution network (MFEN) by way of making MFEB (multi-scale feature extraction blocks) blocks, which step by step obtain multi-scale and hierarchical information. To resolve the issues of network depth as well as width, Qin et al. proposed an ARRFN (Attentive Residual Refinement Network) [48] method. Generally, the architecture of ARRFN consists of feature extraction, multi-scale separable upsampling blocks and attentive residual refinement. Li et al. proposed an adjustable SR network (ASRN) [49], which easily adjusts the network depth of the proposed ASRN model.

3. Proposed method

In this section, we discuss a detailed explanation of our proposed network architecture for SISR known as Squeeze-and-ExcitationNext for Single Image Super-Resolution (SENext), as shown in Fig. 2. The proposed framework mainly consists of two routes/paths with four different types of blocks such as shallow feature extraction block (SFEB), squeeze-and-excitation block (SEB), split-concatenate block (SCB), and finally capsule unit block (CUB) with the support of special upper branch block (UBB). The information transfer pathway passes low, mid, and high-frequency information from the original low-resolution images. In this strategy, we do not change the size of the input image. We extract the features information from the original LR input image and finally add them and pass through the SCB followed by the CUB block. To reconstruct the visually pleasing SHR output, we supply all feature information with a special upper branch, and then the resultant output passes through the learning-based Transpose Convolution layer.

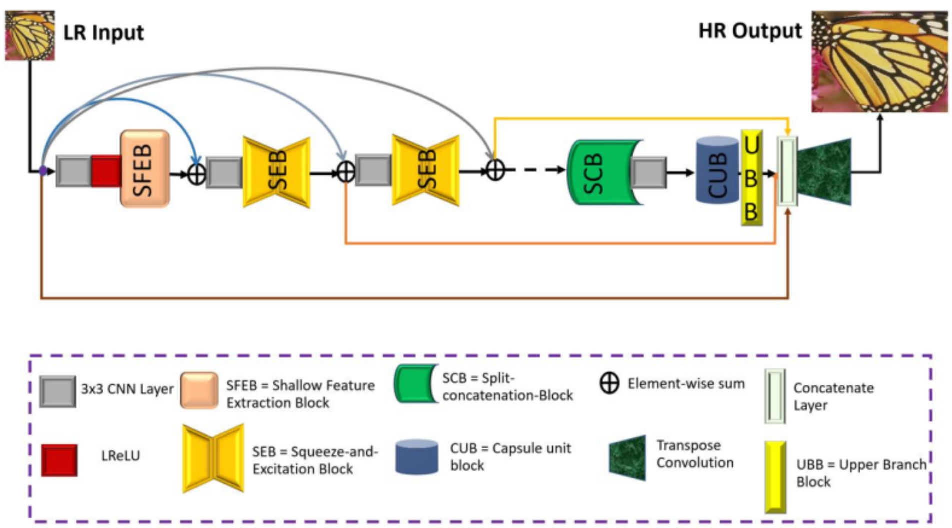


Fig. 2. The proposed framework of Squeeze-and-ExcitationNext for Single Image Super-Resolution (SENext).

3.1. Shallow Feature Extraction Block

According to the survey of [26, 50] the shallow feature F_0 is extracted from the original LR input image using only one or two 3×3 convolutional layers followed by the ReLU activation function, as shown in Fig. 3a and 3b. The design of said blocks is straightforward, but it cannot extract the complete shallow features information from the original LR input image. Furthermore, total network architecture depends on the initial shallow feature extractions, and sometimes essential feature information is lost when a network architecture is significantly deeper. To extract the complete low and high-level features information from the original LR input image, we used the improved version of Fig. 3b, architecture with the use of local (S_L) and global skip (S_G) connections as shown in Fig. 3c. Our proposed, designed shallow feature extraction block is explained as:

$$x_0 = H_{SFEB}(x_{LR\downarrow}),$$

where $H_{SFEB}(.)$ represents convolution operation, and $x_{LR\downarrow}$ is the original input LR image. After obtaining the shallow features x_0 is then used as the input of SEB.

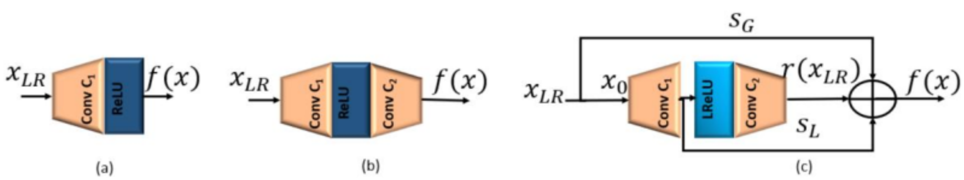


Fig. 3. Different types of Shallow Feature extraction blocks are (a) Single Layer Shallow Feature Extraction blocks (b) Two-layer Shallow Feature Extraction blocks, and (c) Our Proposed Shallow Feature Extraction blocks (SFEB).

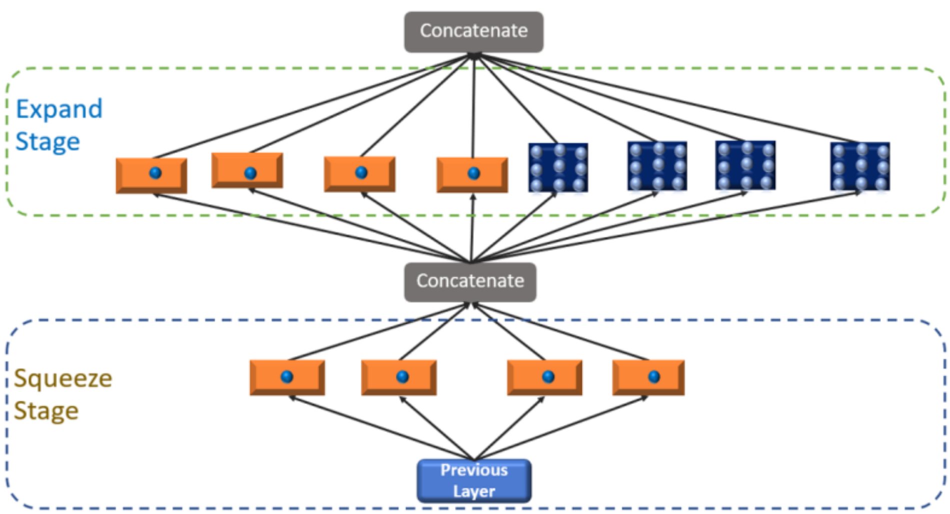


Fig. 4. Original Fire Block (Squeeze and Expand Stage Block).

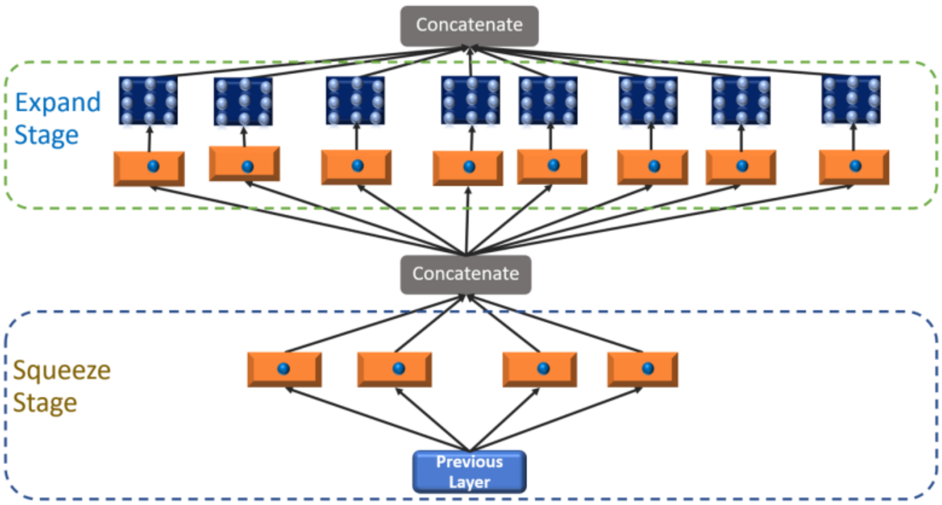


Fig. 5. The proposed fire module is used as a squeeze and excitation (SEB) block.

3.2. SEB BLOCK

For image and computer vision-based applications, the SqueezeNet deep CNN architecture mainly focuses on computational cost and model efficiency [51]. The first basic architecture of the SqueezeNet block is commonly known as a fire module, as shown in Fig. 4. The whole architecture consists of two stages: a squeeze stage that applies a series of 1 × 1 kernel and the expanded stage use 3 × 3 kernels both followed by a conventional rectified linear unit (ReLU) activation function. The number of squeeze filters that can be learned is always less than the volume of the input. Consequently, the squeeze stage may be considered a dimensionality reduction process that also captures the pixel correlations between input channels. The output of the squeezing phase relates to the expansion phase, which combines learning 1 × 1 and 3 × 3 convolutions. To reduce the vanishing gradient issue during the training as well as decrease the computational complexity, we proposed an improved squeeze-and-excitation block (SEB) by stacking a series of 1 × 1 convolution layers in each phase and using the LReLU activation function in place of the ReLU activation function. Suppose the proposed SEB contains N number of Blocks, then x_{n-1} and x_n be the input and output of the nth SEB block. The resultant output of x_n feed to the SCB block.

$$x_n = H_n(x_{n-1}), \\ = H_n(x_{n-1}) \; (\dots(H_1(x_0))\dots)), \qquad (2)$$

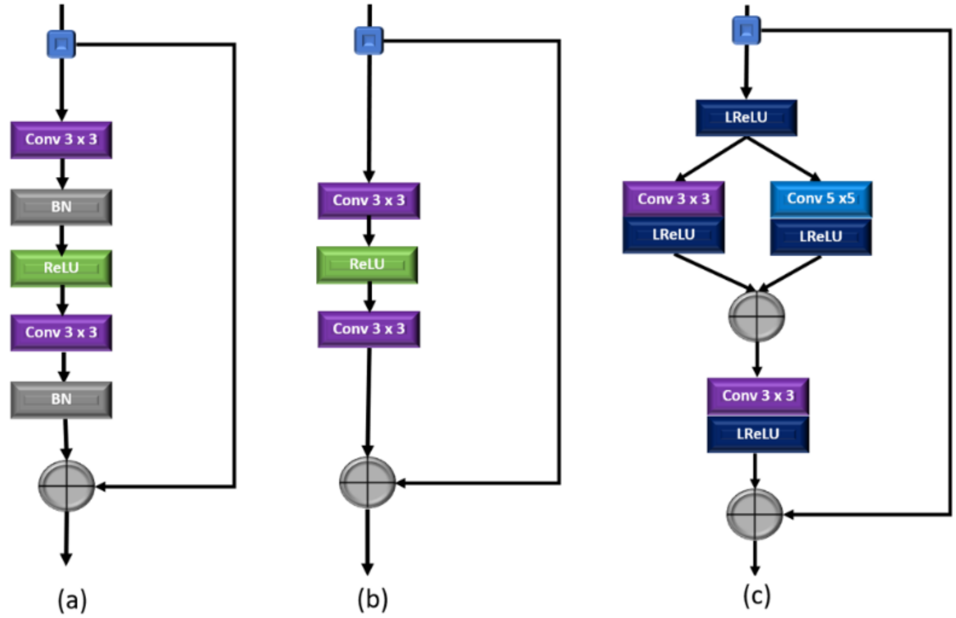


Fig. 6. The structures of different residual learning blocks. (a) SRResNet [25], (b) EDSR [26], and (c) Our proposed Split-Concatenate Block (SCB).

3.3. SCB Block

Residual learning is one of the most crucial methods to make training large-scale networks easier [52]. A global skip connection was implemented by Kim et al. in [9] and could concentrate on predicting the residual skip connection learning. ResNet is a fundamental component of CNN and was created in [52] by applying residual learning to a few stacked layers. More extracted feature information is readily moved through every block using the short-term skip connection [53]. Numerous efforts have altered the structure of ResNet, which was first developed for the image recognition task; its performance has been improved. Several versions of the residual learning-based construction blocks are shown in Figs. 6a and 6b. The SRResNet building block [25] differs from ResNet in lacking the activation layer following element-wise addition. The two batch normalization layers (BN) were eliminated to create the EDSR building blocks when it was suggested [26] that batch normalization (BN) would not be appropriate for the image super-resolution task. Thus, our proposed model adopts a split-concatenate block without BN, as shown in Fig. 6c. Initially, HR features are split into two branches with a kernel size of 3×3 and 5×5 to take the benefit of small as well as a sizeable receptive field both followed by another 3×3 filter with LReLU activation function to prevents gradients from saturating and mitigates the risk of vanishing gradients.

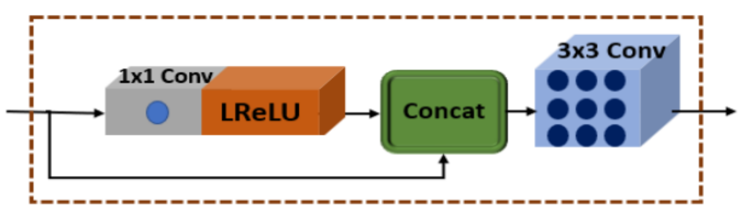


Fig. 7. Proposed Capsule Unit Block (CUB) with local skip connection.

3.4. CUB Block

To minimize the feature map dimension and merge long-term features with skip connections to rebuild the high-quality HR image, a capsule unit is introduced [54]. To follow the concept of [54], we proposed a particular capsule unit block (CUB) with a global skip connection, as shown in Fig. 7. The design of the proposed CUB block consists of one bottleneck layer and one 3 × 3 filter. The bottleneck layer recalibrates the information with a sub-local skip connection to overcome parameter growth and build an efficient architecture. The concatenated output is used by one convolution layer of filter size is 3 × 3.

3.5. UBB Block

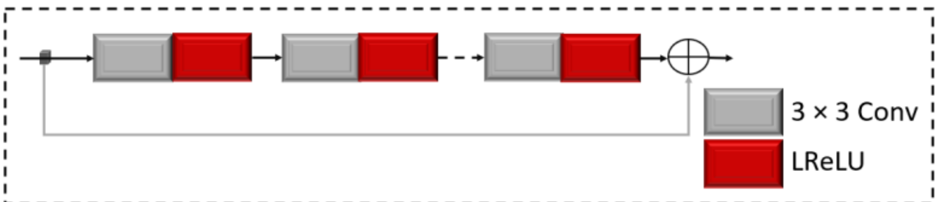


Fig. 8. Proposed Upper Branch Block (UBB) with global skip connection.

Implementing Inception [55] based block before the transpose layer to extract the multi-scale features information obtained the better performance. The main drawback of Inception-based architecture used before the transpose layer is to availability of a max pooling layer. Max pooling layer is to lose the features information, which leads to drop the performance of the model [56, 57]. Furthermore, 5 × 5 kernel size is more time consuming, taking high computational cost and more expensive. To resolve these issues, we proposed an alternate design with a simple upper branch block (UBB) with small kernel size. We removed the max pooling layer operation with a residual skip connection. In the UBB block, we utilized 10 CNN layers having a filter size is 3 × 3 with the support of LReLU function except the last layer.

4. *Experimental Results*

In this section, we assess the effectiveness of our SENext model on different public datasets. Initially, we discuss the training and testing datasets; then, we will explain the experimental evaluations with state-of-the-art methods. Our model training was performed on the combination of two datasets, such as DIV2K [27] (select 100 images of 2K resolution), and BSDS300 [58]. The same combination is observed in [50, 59]. We apply the data augmentation technique to reduce the chances of overfitting and improve training efficiency. For experimental calculations we used the five benchmark test datasets, such as Set5 [60] , Set14 [61] , BSDS100 [58], Urban100 [62] and Manga109 [63]. The original low-resolution image is obtained using MATLAB bicubic operation for enlargement scale factor 2×, 3×, 4×, and 8×. For training purposes, we used Adam optimizer, with an initial learning rate of 0.0001. The proposed methods used the Windows 11 operating system having one GPU (GeForce NVIDIA RTX 2070 GPU) and an Intel(R) Core (TM) i7-9750H CPU @ 2.60GHz having 16.0 GB RAM system. The training and testing phase is performed under Keras 2.6.0 with TensorFlow 2.6.0 environment.

4.1. Quantitative comparisons with existing state-of-the-art-methods

In this paper, we compare our SENext quantitatively with fourteen state-of-the-art methods, such as Bicubic, SRCNN [6], FSRCNN [7], VDSR [9], DRCN [11], LapSRN [21], DRRN [24], MemNet [28], ASRN [49], IDN [40], SRMDNF [34], MFEN_S [47], CARN [32], and IMDN [40]. Table 1 summarizes quantitative results on the five benchmark testing datasets.

Table 1. Presents the quantitative assessment of image SR methods with our SENext. The reported quantitative results of average values of PSNR/SSIM with upscale factors 2×, 3×, 4×, and 8×. Red color with bold quantitative values is recorded as a best value. The blue color with underlined quantitative values is indicated as the 2nd best value.

Method	Factor	≠Params↓	Set5 [60] PSNR↑/SSIM↑	Set14 [61] PSNR↑/SSIM↑	BSDS100 [58] PSNR↑/SSIM↑	Urban100 [62] PSNR↑/SSIM↑	Manga109 [63] PSNR↑/SSIM↑	Average PSNR↑/SSIM↑
Bicubic	2×	-/-	33.66 / 0.9299	30.24 / 0.8688	29.56 / 0.8431	26.88 / 0.8403	30.80 / 0.9339	30.23 / 0.8832
SRCNN[6]	2×	57K	36.66 / 0.9542	32.45 / 0.9067	31.36 / 0.8879	29.50 / 0.8946	35.60 / 0.9663	33.11 / 0.9219
FSRCNN[7]	2×	12K	37.05 / 0.9560	32.66 / 0.9090	31.53 / 0.8920	29.88 / 0.9020	36.67 / 0.9710	33.56 / 0.9260
VDSR[9]	2×	665K	37.53 / 0.9590	33.05 / 0.9130	31.90 / 0.8960	30.77 / 0.9140	37.22 / 0.9750	33.24 / 0.9314
DRCN[11]	2×	1,774K	37.63 / 0.9588	33.04 / 0.9124	31.85 / 0.8942	30.75 / 0.9133	37.55 / 0.9732	34.16 / 0.9304
LapSRN[21]	2×	813K	37.52 / 0.9591	33.08 / 0.9130	31.08 / 0.8950	30.41 / 0.9101	37.27 / 0.9740	33.87 / 0.9302
DRRN[24]	2×	297K	37.74 / 0.9591	33.23 / 0.9136	32.05 / 0.8973	31.23 / 0.9188	37.60 / 0.9736	34.37 / 0.9325
MemNet [28]	2×	677K	37.78 / 0.9597	33.28 / 0.9142	32.08 / 0.8978	31.31 / 0.9195	37.72 / 0.9740	34.43 / 0.9330
ASRN [49]	2×	227K	37.67 / 0.9594	33.19 / 0.9144	31.95 / 0.8970	31.20 / 0.9186	37.79 / 0.9753	34.36 / 0.9329
IDN [40]	2×	553K	37.83 / 0.9600	33.30 / 0.9148	32.08 / 0.8985	31.27 / 0.9196	38.01 / 0.9749	34.50 / 0.9336
SRMDNF [34]	2×	1,511K	37.79 / 0.9601	33.32 / 0.9159	32.05 / 0.8985	31.33 / 0.9204	38.07 / 0.9761	34.51 / 0.9342
MFEN_S [47]	2×	755K	38.03 / 0.9606	33.55 / 0.9171	32.19 / **0.9283**	32.19 / 0.8997	38.77 / 0.9772	34.95 / 0.9366
CARN [32]	2×	1,592K	37.76 / 0.9590	33.52 / 0.9166	32.09 / 0.8978	31.92 / 0.9266	38.36 / 0.9765	34.73 / 0.9353
IMDN [40]	2×	694K	38.00 / 0.9605	33.63 / 0.9177	32.19 / 0.8996	32.17 / 0.9283	**38.88 / 0.9784**	34.97 / **0.9369**
SENext (Our)	2×	97K	**38.04 / 0.9608**	**34.24 / 0.9181**	**32.21** / 0.8997	**32.43 / 0.9287**	38.79 / 0.9774	**35.14 / 0.9369**
Bicubic	3×	-/-	30.39 / 0.8682	27.55 / 0.7742	27.21 / 0.7385	24.46 / 0.7349	26.95 / 0.8566	27.31 / 0.7945
SRCNN[6]	3×	57K	32.75 / 0.9090	29.30 / 0.8215	28.41 / 0.7863	26.24 / 0.7989	30.48 / 0.9117	29.44 / 0.8455
FSRCNN[7]	3×	12K	33.18 / 0.9140	29.37 / 0.8240	28.53 / 0.7910	26.34 / 0.8080	31.10 / 0.9210	29.70 / 0.8516
VDSR[9]	3×	665K	33.66 / 0.9213	29.77 / 0.8314	28.82 / 0.7976	27.14 / 0.8279	32.01 / 0.9340	30.28 / 0.8624
DRCN[11]	3×	1,774K	33.82 / 0.9226	29.76 / 0.8311	28.80 / 0.7963	27.15 / 0.8276	32.24 / 0.9343	30.35 / 0.8624
LapSRN[21]	3×	813K	33.82 / 0.9227	29.87 / 0.8320	28.82 / 0.7980	27.07 / 0.8280	32.21 / 0.9350	30.36 / 0.8631
DRRN[24]	3×	297K	34.03 / 0.9244	29.96 / 0.8349	28.95 / 0.8004	27.53 / 0.8378	32.71 / 0.9379	30.64 / 0.8671
MemNet [28]	3×	677K	34.09 / 0.9248	30.00 / 0.8350	28.96 / 0.8001	27.56 / 0.8376	32.51 / 0.9369	30.62 / 0.8669
ASRN [49]	3×	248K	33.84 / 0.9223	29.97 / 0.8348	28.86 / 0.7990	27.41 / 0.8342	32.63 / 0.9364	30.54 / 0.8653
IDN [40]	3×	553K	34.11 / 0.9253	29.99 / 0.8354	28.95 / 0.8013	27.42 / 0.8359	32.71 / 0.9381	30.64 / 0.8672
SRMDNF [34]	3×	1,528K	34.12 / 0.9254	30.04 / 0.8382	28.97 / 0.8025	27.57 / 0.8398	33.00 / 0.9403	30.74 / 0.8692
CARN [32]	3×	1,592K	34.29 / 0.9255	30.29 / 0.8407	29.06 / 0.8034	28.06 / 0.8493	33.50 / 0.9440	31.04 / 0.8726
IMDN [40]	3×	703K	**34.36 / 0.9270**	30.32 / 0.8417	29.09 / 0.8046	28.17 / **0.8519**	33.61 / 0.9446	31.11 / **0.8740**
SENext (Our)	3×	54K	34.32 / 0.9255	**31.08 / 0.8419**	**29.11 / 0.8047**	**28.60 / 0.8519**	**33.63 / 0.9451**	**31.35** / 0.8738
Bicubic	4×	-/-	28.42 / 0.8104	26.00 / 0.7027	25.96 / 0.6675	23.14 / 0.6577	24.89 / 0.7866	25.68 / 0.7250
SRCNN[6]	4×	57K	30.48 / 0.8628	27.50 / 0.7513	26.90 / 0.7010	24.52 / 0.7221	27.58 / 0.8555	27.40 / 0.7785
FSRCNN[7]	4×	12K	30.72 / 0.8660	27.61 / 0.7550	26.98 / 0.7150	24.62 / 0.7280	27.90 / 0.8610	27.57 / 0.7850
VDSR[9]	4×	665K	31.35 / 0.8838	28.01 / 0.7674	27.29 / 0.7251	25.18 / 0.7524	28.83 / 0.8870	28.13 / 0.8031
DRCN[11]	4×	1,774K	31.53 / 0.8854	28.02 / 0.7670	27.23 / 0.7233	25.14 / 0.7510	28.93 / 0.8854	28.17 / 0.8024
LapSRN[21]	4×	813K	31.54 / 0.8850	28.19 / 0.7720	27.32 / 0.7270	25.21 / 0.7560	29.09 / 0.8900	28.27 / 0.8060
DRRN[24]	4×	297K	31.68 / 0.8888	28.21 / 0.7720	27.38 / 0.7284	25.44 / 0.7638	29.45 / 0.8946	28.43 / 0.8095
MemNet [28]	4×	677K	31.74 / 0.8893	28.26 / 0.7723	27.40 / 0.7281	25.50 / 0.7630	29.42 / 0.8942	28.46 / 0.8094
ASRN [49]	4×	244K	31.65 / 0.8867	28.28 / 0.7733	27.34 / 0.7279	25.42 / 0.7616	29.59 / 0.8935	28.46 / 0.8086
IDN [40]	4×	553K	31.82 / 0.8903	28.25 / 0.7730	27.41 / 0.7297	25.41 / 0.7632	29.41 / 0.8942	28.46 / 0.8101
SRMDNF [34]	4×	1,552K	31.96 / 0.8925	28.35 / 0.7787	27.49 / 0.7337	25.68 / 0.7731	30.09 / 0.9024	28.71 / 0.8161
MFEN_S [47]	4×	775K	32.23 / 0.8951	28.61 / 0.7814	26.07 / 0.7847	27.56 / 0.7355	30.41 / 0.9074	28.98 / **0.8208**
CARN [32]	4×	1,592K	32.13 / 0.8937	28.60 / 0.7806	27.58 / 0.7349	26.07 / 0.7837	30.47 / **0.9084**	28.97 / 0.8203
IMDN [40]	4×	715K	**32.21 / 0.8948**	28.58 / 0.7811	27.56 / 0.7353	26.04 / 0.7838	30.45 / 0.9075	28.97 / 0.8205
SENext (Our)	4×	54K	31.50 / 0.8947	**28.99 / 0.7812**	**28.49 / 0.7357**	**26.64 / 0.7839**	**30.48 / 0.9084**	**29.22 / 0.8208**
Bicubic	8×	-/-	24.40 / 0.6580	23.10 / 0.5660	23.67 / 0.5480	20.74 / 0.5160	21.47 / 0.6500	22.68 / 0.5876
SRCNN[6]	8×	57K	25.34 / 0.6471	23.86 / 0.5443	24.14 / 0.5043	21.29 / 0.5133	22.46 / 0.6606	23.42 / 0.5739
FSRCNN[7]	8×	12K	25.42 / 0.6440	23.94 / 0.5482	24.21 / 0.5112	21.32 / 0.5090	22.39 / 0.6357	23.46 / 0.5696
VDSR[9]	8×	665K	25.73 / 0.6743	23.20 / 0.5110	24.34 / 0.5169	21.48 / 0.5289	22.73 / 0.6688	23.50 / 0.5800
DRCN[11]	8×	1,774K	25.93 / 0.6743	24.25 / 0.5510	24.49 / 0.5168	21.71 / 0.5289	23.20 / 0.6686	23.92 / 0.5879
LapSRN[21]	8×	813K	26.15 / 0.7028	24.45 / 0.5792	24.54 / 0.5293	21.81 / 0.5555	23.39 / 0.7068	24.07 / 0.6147

SENext (Our)	8×	97K	**26.87 / 0.7415**	**25.73 / 0.6200**	**26.79 / 0.5847**	**21.90 / 0.5829**	**23.96 / 0.7389**	**25.05 / 0.6536**

4.2. Comparison analysis based on the number of model parameters

We evaluate the computational cost of our SENext model in terms of the size of the network parameters versus PSNR, as shown in Fig. 9. By employing the squeeze-and-excitation blocks, our SENext network model shrink size of the model in terms of K parameters with other deep CNN image SR methods. The proposed model evaluates the performance on Set5 [60] test dataset with an enlargement scale factor 2×. Our SENext have number of parameters about 85% lower than the VDSR [9], 95% lower than the DRCN [11] [11], 88% lower than the LapSRN [21], 667% lower than the DRRN [24], 86% lower than the MemNet [28], 57% lower than the ASRN [49], 82% lower than the IDN [40], 94% lower than the SRMDNF [34], 87% lower than the MFEN_S [47], 94% lower than the CARN [32], 86% lower than the IMDN [40].

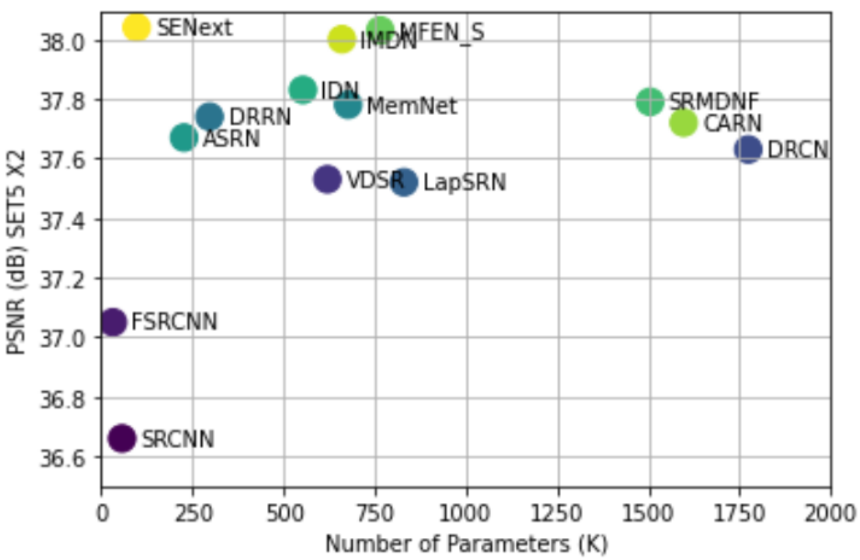


Fig. 9. The performance comparison in terms of model parameters versus PSNR tested on image dataset of Set5 with upscale factor 2×.

4.3. Comparison analysis based on the Image Quality Metrics

In this sub-section, we present the quantitative evaluation in terms of PSNR/SSIM, as shown in Fig. 10. The results demonstrate that our SENext attains the best quantitative performance of existing deep CNN image SR methods. Using a squeeze-and-excitation block with local and global skip connection, our proposed model has obtained the peak value in both quality metrics (PSNR/SSIM).

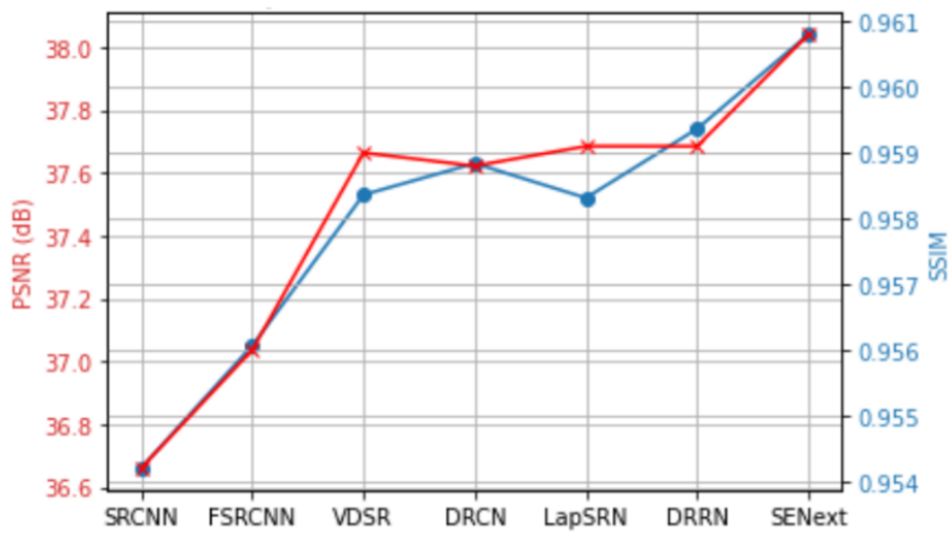


Fig. 10. Quantitative evaluation of average PSNR and SSIM on all test datasets having an enlargement factor 2×.

4.4. Quantitative Analysis of run time versus PSNR

In this section, we assess our SENext model's performance in terms of runtime time versus PSNR, as seen in Fig. 11. To assess the state-of-the-art approaches using a Intel i7-9750H CPU @ 2.60GHz NVIDIA GeForce RTX 2070 GPU (16 GB Memory). For evaluation purposes, we used the public access codes provided by the authors. We use the authors' public access codes for evaluation purposes. The trade-off between CPU time of execution versus PSNR on Set5 [60] enlargement factor 2× is present in Fig. 11. Our proposed method is faster than all state-of-the-art methods except the shallow models (SRCNN and FSRCNN). Furthermore, our proposed SENext attains less computation cost regarding floating-point operations per second (FLOPs), as shown in Fig. 12.

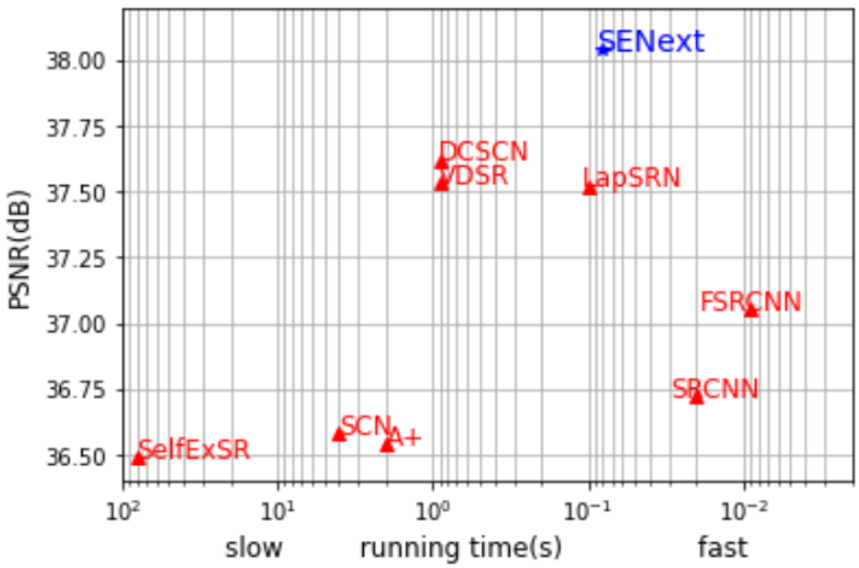


Fig. 11. Running time and accuracy trade-off. The results are evaluated on Set5 with a scale factor of ×2.

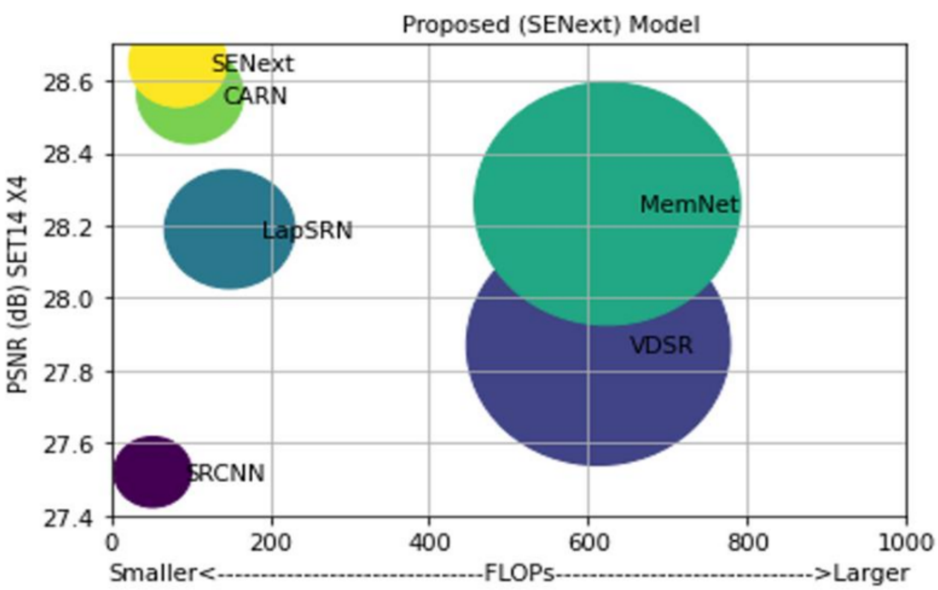
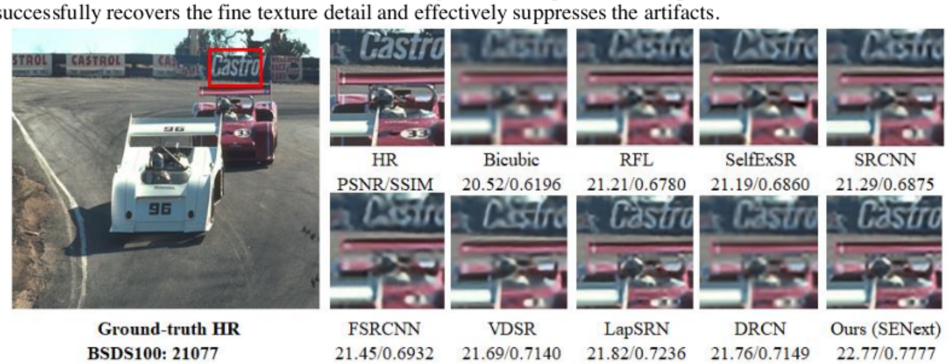
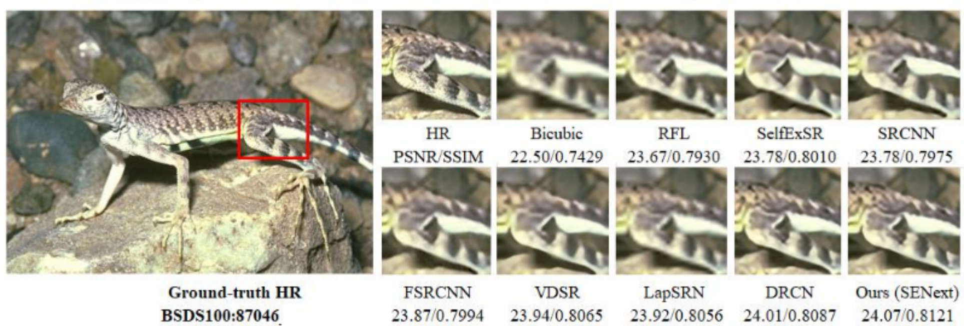


Fig. 12. Quantitative evaluations of PSNR versus FLOPs on Set14 enlargement factor 2×.

4.5. Perceptual Quality Comparison

Fig. 13 and 14 shows the perceptual quality of enlargement factor 4× and 8× image SR test datasets including BSDS100 [58], Urban100 [62] and Manga109 [63]. The results on challenging enlargement scale factor 8× results observed that more blurry results were generated by Bicubic, RFL [5], SelfExSR [62], SRCNN [6], and FSRCNN [7]. However, it is a difficult effort to improve an image for an enlargement factor of 8×, our SENext successfully recovers the fine texture detail and effectively suppresses the artifacts.





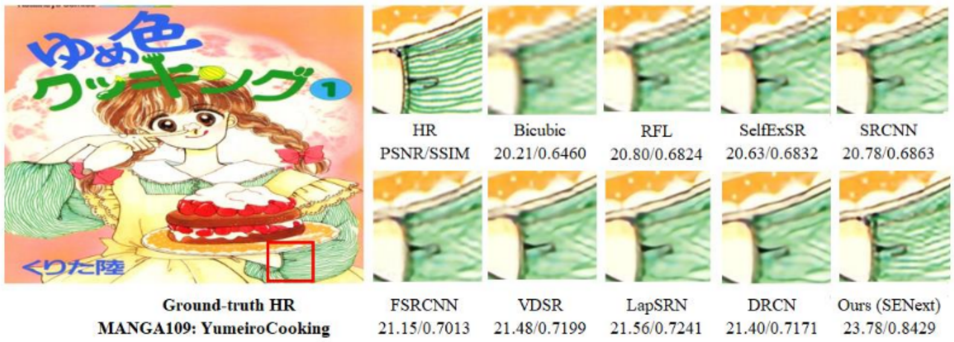


Fig. 13. Visual perceptual quality-wise improvement of different images with 4× enlargement factor on BSDS100, URBAN100, and MANGA109 image datasets.

Fig. 14. Visual perceptual quality-wise improvement of different images with 8× enlargement factor on URBAN100~~,~~ and MANGA109 image datasets.

4.6. Ablation studies

4.6.1. Model Analysis with different Block arrangements.

A more comprehensive ablation study of our proposed blocks can be found in Table 2. In this experiment, we investigated the effects of various combinations of blocks. The eight networks were trained for spatial super-resolution application with enlargement factor 8× and have the same configuration of training as well as validation parameters. We used the 100 images of the DIV2K [27] dataset for training and Yang91 [1] images for validation with 16 batch sizes having 100 epochs. In Table 2 PSNR value is reported and observed that the baseline network (without any block) gives the lowest PSNR value (28.11 dB), but the best performance (28.48 dB) is observed when all blocks are used in the model.

Table 2. Ablation study of different blocks, including SFEB, SEB, and SCB. The quantitative value of average PSNR calculated on Set14 enlargement factor 4x on 100 epochs.

Blocks	Combination of Blocks							
SFEB	□	✓	□	✓	□	✓	□	✓
SEB	□	□	✓	✓	□	□	✓	✓
SCB	□	□	□	□	✓	✓	✓	✓
Average PSNR	28.11	28.23	28.20	28.35	28.38	28.42	28.45	28.48

4.6.2. Selection of Optimizers.

The selection of an optimizer plays a crucial role during the training to optimize the model efficiency and reduce the chance of overfitting. Our proposed SENext model is trained on four different optimizers, including Adam [64], Adamax, which is an enhanced version of Adam, and stochastic gradient descent (SGD). The experimental results with loss function as shown in Fig 15. A more stable pattern of Adam appears in Fig. 15. In the case of RMSprop (green line) decreases slowly with more ripples after 400 iterations as compared to Adam. All optimizers were trained on 400 epochs with the base model. We used the 100 images of the DIV2K [27] dataset for training and Yang91 [1] images for validation with 16 batch sizes.

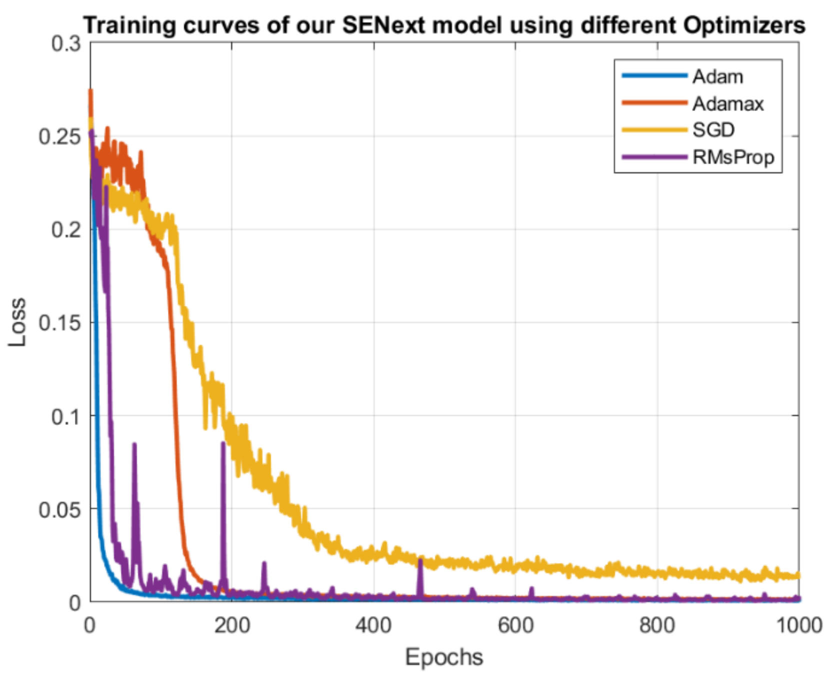


Fig. 15. Training curves optimization with different optimizers.

5. *Conclusion and Future work*

In this study proposes a novel two-stage squeeze (compress) and expand network architecture for single image super-resolution (SENext). Proposed SENext used SFEB, SEB, SCB, CUB, and UBB blocks with the support of local and global skip connections. The SFEB block extracts the low-frequency features from the original LR image. The resultant features are fed to the remaining blocks through a long and short skip route. Implementation of SEB side-by-side reduces the computational cost of the model and calculates the high-frequency features information. The use of extensive sub-local skip connections help to reduce vanishing gradient problems during the training. In addition, to activate the dead neurons in the model during the training, we replaced the conventional ReLU activation function with LReLU. Furthermore, the comparative analysis and ablation study shows the efficiency of a squeeze and excitation network to reduce lots of parameters and computations only with slight performance drops. Extensive evaluations on five benchmark test datasets show that using a large upsampling factor of 4× or 8× improves the reconstruction results in both quantitative and qualitative criteria. In the future, we will further optimize our model to introduce multi-path learning with dense global and local skip connections under complex scenarios. Future work will involve further model optimization to implement multi-path learning with dense global and local skip connections under complex scenarios.

Author contributions

"This manuscript was performed in collaboration between the authors. Wazir Muhammad proposed the new SISR method based on squeeze-and-excitation blocks. Wazir Muhammad, Supavadee Aramvith, and Takao Onoye were involved in the writing and reviewing of the manuscript. All authors discussed and approved the final manuscript for final submission".

CRediTauthorship contribution statement

Wazir Muhammad: Conceptualization, Methodology, Software, Validation, Writing - original draft. **Supavadee Aramvith**: Methodology, Supervision, Writing - review & editing. **Takao Onoye**: Writing - review & editing.

Declaration of competing interest

"The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper".

Acknowledgements

"This work is supported by the Second Century Fund (C2F) Chulalongkorn University Bangkok Thailand, Electrical Engineering Department Chulalongkorn University Bangkok, Thailand, Thailand Science research and Innovation Fund Chulalongkorn University (CU_FRB65_ind (9)_157_21_23), (CU_FRB65_soc-(1)_001_27_01), Ratchadaphiseksomphot Endowment Fund (Multimedia Data Analytics and Processing Research Unit)".

The authors thank the editors and reviewers for their valuable comments and suggestions.

References

1. Yang, J., et al., *Image super-resolution via sparse representation.* IEEE transactions on image processing, 2010. **19**(11): p. 2861-2873.

2. Yang, J., et al. *Image super-resolution as sparse representation of raw image patches*. in *2008 IEEE conference on computer vision and pattern recognition*. 2008. IEEE.
3. Yang, C.-Y. and M.-H. Yang. *Fast direct super-resolution by simple functions*. in *Proceedings of the IEEE international conference on computer vision*. 2013.
4. Timofte, R., V. De Smet, and L. Van Gool. *A+: Adjusted anchored neighborhood regression for fast super-resolution*. in *Asian conference on computer vision*. 2014. Springer.
5. Schulter, S., C. Leistner, and H. Bischof. *Fast and accurate image upscaling with super-resolution forests*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.
6. Dong, C., et al., *Image super-resolution using deep convolutional networks.* 2015. **38**(2): p. 295-307.
7. Dong, C., C.C. Loy, and X. Tang. *Accelerating the super-resolution convolutional neural network*. in *European conference on computer vision*. 2016. Springer.
8. Krizhevsky, A., I. Sutskever, and G.E. Hinton, *Imagenet classification with deep convolutional neural networks.* Advances in neural information processing systems, 2012. **25**.
9. Kim, J., J. Kwon Lee, and K. Mu Lee. *Accurate image super-resolution using very deep convolutional networks*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
10. Shi, W., et al. *Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
11. Kim, J., J. Kwon Lee, and K. Mu Lee. *Deeply-recursive convolutional network for image super-resolution*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
12. Anwar, S., S. Khan, and N. Barnes, *A deep journey into super-resolution: A survey.* ACM Computing Surveys (CSUR), 2020. **53**(3): p. 1-34.
13. Hu, J., L. Shen, and G. Sun. *Squeeze-and-excitation networks*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
14. Cheng, X., et al. *SESR: Single image super resolution with recursive squeeze and excitation networks.* in *2018 24th International conference on pattern recognition (ICPR)*. 2018. IEEE.
15. Maas, A.L., A.Y. Hannun, and A.Y. Ng. *Rectifier nonlinearities improve neural network acoustic models*. in *Proc. icml*. 2013. Citeseer.
16. He, K., et al. *Delving deep into rectifiers: Surpassing human-level performance on imagenet classification*. in *Proceedings of the IEEE international conference on computer vision*. 2015.
17. Wang, Z., et al. *Deep networks for image super-resolution with sparse prior*. in *Proceedings of the IEEE international conference on computer vision*. 2015.
18. Wang, Z., et al., *Deeply improved sparse coding for image super-resolution.* 2015. **2**(3): p. 4.
19. Mao, X., C. Shen, and Y.-B. Yang. *Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections*. in *Advances in neural information processing systems*. 2016.
20. Romano, Y., J. Isidoro, and P. Milanfar, *RAISR: rapid and accurate image super resolution.* IEEE Transactions on Computational Imaging, 2016. **3**(1): p. 110-125.

21. Lai, W.-S., et al. *Deep laplacian pyramid networks for fast and accurate super-resolution*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
22. Zhang, K., et al., *Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising*. IEEE Transactions on Image Processing, 2017. **26**(7): p. 3142-3155.
23. Zhao, Y., et al., *GUN: Gradual upsampling network for single image super-resolution.* IEEE Access, 2018. **6**: p. 39363-39374.
24. Tai, Y., J. Yang, and X. Liu. *Image super-resolution via deep recursive residual network*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
25. Ledig, C., et al., *Photo-realistic single image super-resolution using a generative adversarial network.* arXiv preprint, 2017.
26. Lim, B., et al. *Enhanced deep residual networks for single image super-resolution*. in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2017.
27. Timofte, R., et al. *Ntire 2017 challenge on single image super-resolution: Methods and results*. in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2017.
28. Tai, Y., et al. *Memnet: A persistent memory network for image restoration*. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017.
29. Yamanaka, J., S. Kuwashima, and T. Kurita. *Fast and accurate image super resolution by deep CNN with skip connection and network in network*. in *Neural Information Processing*. 2017. Springer.
30. Han, W., et al. *Image super-resolution via dual-state recurrent networks*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
31. Li, J., et al. *Multi-scale residual network for image super-resolution*. in *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018.
32. Ahn, N., B. Kang, and K.-A. Sohn. *Fast, accurate, and lightweight super-resolution with cascading residual network*. in *Proceedings of the European conference on computer vision (ECCV)*. 2018.
33. Choi, J.-H., et al., *Lightweight and Efficient Image Super-Resolution with Block State-based Recursive Network.* 2018.
34. Zhang, K., W. Zuo, and L. Zhang. *Learning a single convolutional super-resolution network for multiple degradations*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
35. Muhammad, W. and S.J.E. Aramvith, *Multi-Scale Inception Based Super-Resolution Using Deep Learning Approach.* 2019. **8**(8): p. 892.
36. Wang, R., M. Gong, and D.J.I.T.o.I.P. Tao, *Receptive Field Size Versus Model Depth for Single Image Super-Resolution*. 2019. **29**: p. 1669-1682.
37. Wang, Y., et al., *End-to-end image super-resolution via deep and shallow convolutional networks.* 2019. **7**: p. 31959-31970.
38. Yang, X., et al., *DRFN: Deep Recurrent Fusion Network for Single-Image Super-Resolution With Large Factors.* 2019. **21**(2): p. 328-337.
39. Su, M., et al. *Hierarchical Recursive Network for Single Image Super Resolution*. in *2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. 2019. IEEE.

40. Hui, Z., et al. *Lightweight image super-resolution with information multi-distillation network*. in *Proceedings of the 27th acm international conference on multimedia*. 2019.
41. Hung, K.-W., Z. Zhang, and J.J.I.A. Jiang, *Real-time image super-resolution using recursive depthwise separable convolution network.* 2019. **7**: p. 99804-99816.
42. Barzegar, S., et al., *Super-resolution using lightweight detailnet network.* 2020. **79**(1): p. 1119-1136.
43. Muhammad, W., S. Aramvith, and T. Onoye, *Multi-scale Xception based depthwise separable convolution for single image super-resolution.* Plos one, 2021. **16**(8): p. e0249278.
44. Hsu, J.-T., C.-H. Kuo, and D.-W.J.I.A. Chen, *Image Super-Resolution Using Capsule Neural Networks.* 2020. **8**: p. 9751-9759.
45. Liu, B. and D.J.N. Ait-Boudaoud, *Effective image super resolution via hierarchical convolutional neural network.* 2020. **374**: p. 109-116.
46. Yang, A., et al., *Non-linear perceptual multi-scale network for single image super-resolution.* Neural Networks, 2022. **152**: p. 201-211.
47. Xiao, H., et al., *MFEN: Lightweight multi-scale feature extraction super-resolution network in embedded system.* Microprocessors and Microsystems, 2022: p. 104568.
48. Qin, J. and R. Zhang, *Lightweight Single Image Super-Resolution with Attentive Residual Refinement Network.* Neurocomputing, 2022.
49. Li, J., et al., *Adjustable Super-Resolution Network via Deep Supervised Learning and Progressive Self-Distillation*. Neurocomputing, 2022.
50. Zhang, Y., et al. *Image super-resolution using very deep residual channel attention networks*. in *Proceedings of the European conference on computer vision (ECCV)*. 2018.
51. Iandola, F.N., et al., *SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size*. arXiv preprint arXiv:1602.07360, 2016.
52. He, K., et al. *Deep residual learning for image recognition*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
53. Chang, K., et al., *Accurate single image super-resolution using multi-path wide-activated residual network.* Signal Processing, 2020. **172**: p. 107567.
54. Zhou, Y., et al., *Image super-resolution based on dense convolutional auto-encoder blocks.* Neurocomputing, 2021. **423**: p. 98-109.
55. Szegedy, C., et al. *Going deeper with convolutions*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.
56. Nirthika, R., et al., *Pooling in convolutional neural networks for medical image analysis: a survey and an empirical study.* Neural Computing and Applications, 2022: p. 1-27.
57. Sabour, S., N. Frosst, and G.E. Hinton, *Dynamic routing between capsules.* Advances in neural information processing systems, 2017. **30**.
58. Arbelaez, P., et al., *Contour detection and hierarchical image segmentation.* IEEE transactions on pattern analysis and machine intelligence, 2010. **33**(5): p. 898-916.
59. Li, F., H. Bai, and Y. Zhao, *Detail-preserving image super-resolution via recursively dilated residual network.* Neurocomputing, 2019. **358**: p. 285-293.
60. Bevilacqua, M., et al., *Low-complexity single-image super-resolution based on nonnegative neighbor embedding.* 2012.

61. Zeyde, R., M. Elad, and M. Protter. *On single image scale-up using sparse-representations*. in *International conference on curves and surfaces*. 2010. Springer.
62. Huang, J.-B., A. Singh, and N. Ahuja. *Single image super-resolution from transformed self-exemplars*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.
63. Matsui, Y., et al., *Sketch-based manga retrieval using manga109 dataset.* Multimedia Tools and Applications, 2017. **76**(20): p. 21811-21838.
64. Kingma, D.P. and J. Ba, *Adam: A method for stochastic optimization.* arXiv preprint arXiv:1412.6980, 2014.

SENExt

ORIGINALITY REPORT

17%	13%	14%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.ncbi.nlm.nih.gov Internet Source	3%
2	www.mdpi.com Internet Source	1%
3	"Computer Vision – ECCV 2020 Workshops", Springer Science and Business Media LLC, 2020 Publication	1%
4	dokumen.pub Internet Source	1%
5	deepai.org Internet Source	1%
6	arxiv.org Internet Source	1%
7	link.springer.com Internet Source	1%
8	paper.ijcsns.org Internet Source	1%

9	Wazir Muhammad, Zuhaibuddin Bhutto, Arslan Ansari, Mudasar Latif Memon et al. "Multi-Path Deep CNN with Residual Inception Network for Single Image Super-Resolution", Electronics, 2021 Publication	1 %
10	journals.plos.org Internet Source	1 %
11	discovery.ucl.ac.uk Internet Source	<1 %
12	"Advances in Brain Inspired Cognitive Systems", Springer Science and Business Media LLC, 2018 Publication	<1 %
13	"Computer Vision – ECCV 2018 Workshops", Springer Science and Business Media LLC, 2019 Publication	<1 %
14	coek.info Internet Source	<1 %
15	Kan Chang, Minghong Li, Pak Lun Kevin Ding, Baoxin Li. "Accurate single image super-resolution using multi-path wide-activated residual network", Signal Processing, 2020 Publication	<1 %
16	www.pubfacts.com Internet Source	<1 %

17	www.spiedigitallibrary.org Internet Source	<1 %
18	Hang Xiao, Jiayi Qin, Seunggil Jeon, Binyu Yan, Xiaomin Yang. "MFEN: Lightweight multi-scale feature extraction super-resolution network in embedded system", Microprocessors and Microsystems, 2022 Publication	<1 %
19	awesome-davian.gitbook.io Internet Source	<1 %
20	Submitted to University of Liverpool Student Paper	<1 %
21	Wazir Muhammad, Supavadee Aramvith, Takao Onoye. "Multi-scale Xception based depthwise separable convolution for single image super-resolution", PLOS ONE, 2021 Publication	<1 %
22	www.biorxiv.org Internet Source	<1 %
23	www.jmedicalcasereports.com Internet Source	<1 %
24	Farnaz Garehdaghi, Saeed Meshgini, Reza Afrouzian. "Positron emission tomography image enhancement using magnetic resonance images and U-net structure", Computers & Electrical Engineering, 2021	<1 %

Publication

25	ebin.pub Internet Source	<1 %
26	tohoku.pure.elsevier.com Internet Source	<1 %
27	unsworks.unsw.edu.au Internet Source	<1 %
28	www.jsjkx.com Internet Source	<1 %
29	"Computer Vision – ECCV 2018", Springer Science and Business Media LLC, 2018 Publication	<1 %
30	econpapers.repec.org Internet Source	<1 %
31	www.eurasip.org Internet Source	<1 %
32	"Advances in Multimedia Information Processing – PCM 2018", Springer Science and Business Media LLC, 2018 Publication	<1 %
33	Jae-Seok Choi, Munchurl Kim. "Single Image Super-Resolution Using Global Regression Based on Multiple Local Linear Mappings", IEEE Transactions on Image Processing, 2017 Publication	<1 %

34 Yuan Zhou, Yeda Zhang, Xukai Xie, Sun-Yuan Kung. "Image super-resolution based on dense convolutional auto-encoder blocks", Neurocomputing, 2021 <1 %
Publication

35 www.arxiv-vanity.com <1 %
Internet Source

36 www.researchgate.net <1 %
Internet Source

37 "Advances in Computer Graphics", Springer Science and Business Media LLC, 2019 <1 %
Publication

38 "Artificial Intelligence and Security", Springer Science and Business Media LLC, 2020 <1 %
Publication

39 "Pattern Recognition and Computer Vision", Springer Science and Business Media LLC, 2019 <1 %
Publication

40 Chunwei Tian, Yong Xu, Wangmeng Zuo, Bob Zhang, Lunke Fei, Chia-Wen Lin. "Coarse-to-fine CNN for image super-resolution", IEEE Transactions on Multimedia, 2020 <1 %
Publication

41 Kassiani Zafeirouli, Anastasios Dimou, Apostolos Axenopoulos, Petros Daras. "Efficient, Lightweight, Coordinate-Based <1 %

Network for Image Super Resolution", 2019 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), 2019
Publication

42	Xinrui Jiang, Nannan Wang, Jingwei Xin, Xi Yang, Yi Yu, Xinbo Gao. "Image Super-Resolution via Multi-view Information Fusion Networks", Neurocomputing, 2020 Publication	<1 %
43	Yifan Wang, Lijun Wang, Hongyu Wang, Peihua Li. "End-to-End Image Super-Resolution via Deep and Shallow Convolutional Networks", IEEE Access, 2019 Publication	<1 %
44	backend.orbit.dtu.dk Internet Source	<1 %
45	export.arxiv.org Internet Source	<1 %
46	mdpi-res.com Internet Source	<1 %
47	www.techscience.com Internet Source	<1 %
48	"Image and Graphics", Springer Science and Business Media LLC, 2021 Publication	<1 %
49	"MultiMedia Modeling", Springer Nature, 2017 Publication	<1 %

50	Chunwei Tian, Ruibin Zhuge, Zhihao Wu, Yong Xu, Wangmeng Zuo, Chen Chen, Chia-Wen Lin. "Lightweight image super-resolution with enhanced CNN", Knowledge-Based Systems, 2020 Publication	<1%
51	"MultiMedia Modeling", Springer Science and Business Media LLC, 2018 Publication	<1%
52	Leilei Zhu, Shu Zhan, Haiyan Zhang. "Stacked U-shape networks with channel-wise attention for image super-resolution", Neurocomputing, 2019 Publication	<1%
53	Xinxia Fan, Yanhua Yang, Cheng Deng, Jie Xu, Xinbo Gao. "Compressed multi-scale feature fusion network for single image super-resolution", Signal Processing, 2018 Publication	<1%
54	Xiumei Wang, Yanan Gu, Xinbo Gao, Zheng Hui. "Dual residual attention module network for single image super resolution", Neurocomputing, 2019 Publication	<1%

Exclude quotes Off Exclude matches Off