Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

A Comprehensive Review of Super-Resolution Techniques: Progress and Prospects from SRCNN to ESRGAN

Sandhra Merin Sabu, Dr. Jubilant J Kizhakettotham

*1*Department of Computer Science and Engineering, Saintgits College of Engineering, Pathamuttam, Kottyam, 686532, kerala, India.*
*2*Department of Computer Science and Engineering, Saintgits College of Engineering, Pathamuttam, Kottyam, 686532, kerala, India.*
**Corresponding author(s). E-mail(s): sandhra.se2325@saintgits.org;*
Contributing authors: jubilant.j@saintgits.org;

ARTICLE INFO

Received: 15 Nov 2024

Revised: 27 Dec 2024

Accepted: 10 Jan 2025

ABSTRACT

In computer vision, Super-Resolution (SR) is considered to an essential technology that tackles the problem of reconstructing high-resolution (HR) images from it's low-resolution (LR) inputs. With the backing of developments in deep learning and computational architectures, numerous SR models have been created over time. From early CNN-based techniques like that of SRCNN and VDSR to more sophisticated designs like EDSR, RCAN and the attention-enhanced models like SAN and HAN, a variety of SR models have been created to tackle various challenges. Numerous Real-time and resource-constrained applications might be benefit from lightweight models like CARN and DRRN that balance efficiency and quality, whereas the Generative Adversarial Networks (GANs) such as SRGAN and ESRGAN may prioritize perceptual quality. These models are widely applicable in variety of domains such as gaming, entertainment, satellite imaging, and healthcare. This research tries to outlines the evolution, strengths and applications of SR models highlighting how they can contribute to improving image quality and outlining potential paths for accessibility and optimization in the future.

Keywords: Super-resolution, Super-Resolution Convolutional Neural Network, Enhanced Super-Resolution Generative Adversarial Network, Residual Dense Network, Hierarchical Attention Network, Enhanced Deep Super-Resolution Network, Residual Channel Attention Network

1. INTRODUCTION

Reconstructing or recreating high-resolution (HR) images from their low-resolution (LR) predecessors is the goal of the crucial field of super-resolution (SR) in computer vision. By enhancing clarity and detail, SR gets beyond restrictions brought on by outdated data, bandwidth issues, or image gear. This feature helps to improves the visual quality and has extensive application in digital content restoration, medical imaging, satellite images and video streaming. As Super-resolution techniques have developed they have moved from basic interpolation-based approaches to complex deep learning structures where each of which addresses particular application needs and challenges[1].

In its most fundamental usecase SR solves the problem of providing precise and comprehensive images in situations when it is not feasible to obtain high-resolution data. As an instance, technical limitations or distance may result in low-resolution images from satellite cameras or obsolete imaging equipment. In a similar vein, video streaming services frequently reduce resolution in order to conserve bandwidth, hence SR is crucial for improving in real time and satisfying consumers.

From a technological perspective, there are two types of SR techniques: traditional and modern. Bilinear or bicubic interpolation, which is computationally efficient but had troubles in producing crisp, detailed outputs, were the interpolation techniques used in ancient systems. On the other hand modern SR methods use machine learning especially deep learning, to identify intricate patterns and generate intricate and accurate reconstructions. These techniques range from straightforward Convolutional Neural Networks (CNNs) to that of the most sophisticated architectures that employ the Generative Adversarial Networks (GANs) and approaches for attention.

Interpolation-based approaches includes bilinear, nearest-neighbor and bicubic interpolation were considered to be the foundation of early SR techniques. These techniques use various mathematical procedures to estimate pixel values in order to upscale images. Despite being computationally efficient, they frequently fail to recover fine details and generate blurred images this is because they do not consider the acquired knowledge about the underlying data structure.

A major improvement in image quality was brought by the introduction of the learning-based techniques for SR with the development of deep learning in machine learning. Among the earliest deep learning models for SR were Convolutional Neural Networks (CNNs), including the groundbreaking SRCNN (Super-Resolution CNN). SRCNN uses a simple yet efficient CNN structure to learn a mapping from LR to HR pictures. This was improved upon by its successor, VDSR (Very Deep SR Network), which used deeper topologies to recover finer details and capture complicated patterns, although with the expense of higher processing costs. Advanced architectures like RCAN (Residual Channel Attention Network) and EDSR (Enhanced Deep SR Network) were created as the result of additional developments that has been happened in the domain.

When Generative Adversarial Networks (GANs) were introduced, SR was completely transformed by focusing more emphasis on perceptual quality than pixel-by-pixel precision. The first GAN-based SR model was the SRGAN, which has been optimized for the perceptual loss to generate realistic and aesthetically pleasing outcomes. This was further enhanced by the ESRGAN (Enhanced SRGAN), which produced sharper, artifact-free results by improving the network architecture and loss algorithms. These models function especially well for applications like digital artwork and media restoration where aesthetics are crucial.Additionally, when corresponded LR-HR datasets are not available, unsupervised SR models—like those based on CycleGAN—adopt robust mappings through adversarial training and domain adaptation[2].

Applications for SR are numerous. It helps improve low-resolution scans, such MRIs or CTs, in medical imaging to give greater detail for precise diagnosis. SR enhances the clarity of satellite images, which is important for disaster relief, urban planning, and environmental monitoring[3]. Enhancing footage from low-quality cameras allows for better identification and analysis for security and surveillance purposes.SR is used in the entertainment sector for video upscaling, which enables the conversion of old films or standard-definition material into 4K or high-definition formats. SR is also been used to upscale visuals in real-time in games, improving the visual experience of the user without significantly affecting its performance. Through the reconstruction of intricate features in old or damaged artwork, SR also helps with digital art restoration.

Hence, SR is a significant development in the domain of image processing that has many uses. As technology develops SR keeps up with the times and opening up new avenues for creative solutions in domains that need higher image clarity and resolution[4]. SR is still considered as a vital tool in modern advances in technology, whether it is being used for practical purposes like medical imaging as well as artistic purposes like digital media enhancement.

2. LITERATURE SURVEY

Single-image super-resolution (SISR) has witnessed a significant advancements in transitioning from traditional interpolation methods to the most modern deep learning techniques that may offer high-quality image reconstruction. A three-layer convolutional network was used by the SRCNN[5] to map the low to high-resolution images. Despite being better than conventional techniques it lacked in natural textures and fine details.

ESPCN[6] reduced the computational cost by introducing the sub-pixel convolutions that can increase efficiency by delaying upscaling to the final layers. It produced sharper outputs that are considered to be more in line with human visual perception by substituting the widely used Structural Similarity Index Measure.

A deeper residual network was proposed in VDSR[7] which provides steady training and increased accuracy. Its intricate design frequently produced unnaturally smooth surfaces and required more processing power. In order to improve the visual quality and better capture high-frequency data the RCAN[8] implemented Residual-in-Residual (RIR) structures and Channel Attention (CA) methods.

Second-Order Channel Attention (SOCA) and non-local residual groups were introduced in sophisticated attention models such as the SAN[9] in order to acquire long-range dependencies and the discriminative feature correlations. Through the integration of Layer Attention and Channel-Spatial modules the HAN[10] improved feature refinement at all its hierarchical levels, further expanding this idea.

In order to balance speed and accuracy the CARN[11] explored thin but efficient architectures that can made it

appropriate for the real-time applications. Similarly the RDN[12] improved data flow and preserved resilience across datasets by utilizing Residual Dense Blocks (RDBs) to fully leverage local and global hierarchical characteristics.
In order to increase the receptive fields without raising parameters and the recursive techniques using deep recursive and residual learning were introduced in DRCN[13] and DRRN[14]. High-frequency detail recovery and training stability were improved by these models.
SRGAN[15], which uses adversarial learning and perceptual loss to produce realistic textures, using generator and discriminator networks[16][17]. By leveraging Residual-in-Residual Dense Blocks, avoiding batch normalization and employing relativistic GAN loss for better texture consistency and realism, ESRGAN[18] was developed to address visual artifacts raised in SRGAN.
These varied methods have greatly improved the state-of-the-art in super-resolution, opening the door for high-fidelity, perceptually rich, and computationally effective image enhancement solutions in a variety of fields, including digital restoration, satellite analysis, and medical imaging.

3. DIFFERENT MODELS IN SUPER RESOLUTION

Several models in superresolution have been developed to improve image resolution. The basic superresolution models consist of Input layers, Feature Extraction layers, Feature Processing layers (Residual/Dense Blocks, Attention), Up-sampling, and Reconstruction. The basic architecture is shown in the figure 3.1. These models cover a wide spectrum, from conventional methods to the most advanced deep learning strategies. The most well-known models utilized in superresolution are summarized here:

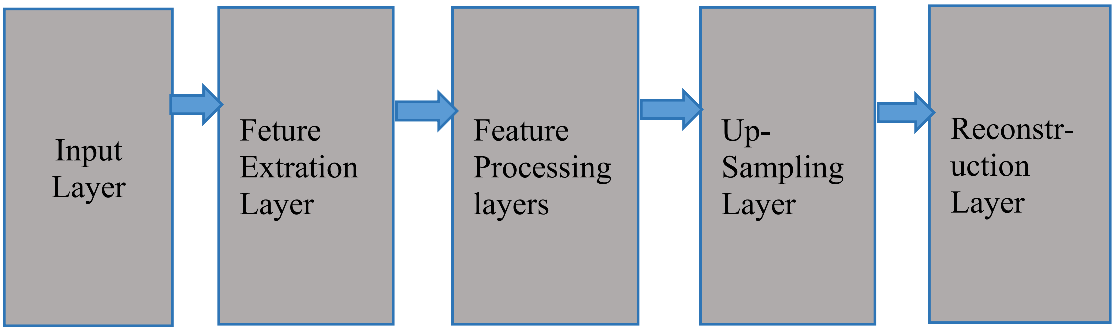

Fig 3.1: Basic architecture of superresolution

a. **SRCNN (Super-Resolution Convolutional Neural Network):**

The Super-Resolution Convolutional Neural Network (SRCNN), which was introduced by Dong et al. in 2014[19] , is among the first deep learning models for single-image superresolution (SISR). SRCNN effectively demonstrated the practical application of convolutional neural networks (CNNs) for image enhancing tasks, laying the groundwork for other advanced superresolution models. With its higher performance in terms of accuracy and image quality, it signified a considerable divergence from conventional interpolation and optimization-based methods.

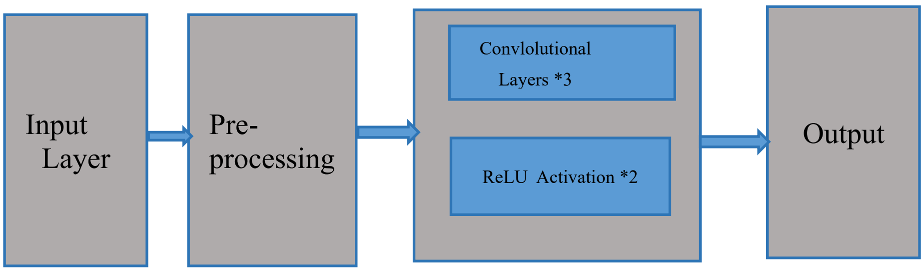

Fig 3.2 : Basic architecture of SRCNN

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

It has a few characteristics, such as being a very simple model with a moderate amount of layers and filters and the maximum level of accuracy. It operates at the fastest possible speed and can improve accuracy through employing an extensive dataset.

It involves several phases, such as preprocessing to increase the image quality using conventional bicubic methods, patch extraction and representation to pass the upscaled images to the first convolutional layer, non-linear mapping to map LR images to HR images, and reconstruction, where features are passed to the third layer to reconstruct the original HR images[20].

Three layers of convolution make up the architecture. Patch extraction is the function of the first layer. Extracting picture patches and identifying fundamental elements like edges and textures are the responsibilities of the first layer. A greater context around each pixel may be captured using the convolution kernel with size of 9×9 . The low-resolution image's patches are extracted (overlapping) via this procedure, which then represents each patch as a high-dimensional vector. The recovered features are subjected to a non-linear transformation by the second layer, which has a kernel size of 1×1, which effectively maps the features to a higher-dimensional space. The mapping between each high-dimensional vector and another high-dimensional vector is nonlinear in this process. The concept of a high-resolution patch is represented by each mapped vector. Reconstructing the HR picture from the mapped feature space is the last layer, which has a kernel size of 5×5. The final high-resolution image is produced by aggregating the aforementioned high-resolution patchwise representations. It is anticipated that this picture will resemble the actual[21].

To add non-linearity between layers, the Rectified Linear Unit (ReLU)[22] is utilized as the activation function. The loss function is Mean Squared Error (MSE), which motivates the network to reduce the pixel-by-pixel discrepancy between the ground truth HR image and the predicted HR image.

Compared to basic interpolation-based methods, it yields results that are crisper and more aesthetically pleasing, but it still has trouble with fine details in highly detailed images, such as textures or edges. SRCNN's computational effectiveness is one of its advantages. It requires little in the way of processing power or massive volumes of data, making it comparatively simple to train as a shallow network.

The learning capacity of SRCNN is restricted by the three-layer CNN, particularly in areas with fine details or complex textures. High-frequency regions become blurry as a result. Additionally, the preprocessing phase of bicubic interpolation used by SRCNN may generate noise or artifacts during the first upscaling of the image. All these affect the quality of the image. CT-SRCNN (Cascade Trained Super Resolution Convolutional Neural Network)[23] is a solution for the above problems. The accuracy may be steadily raised with additional layers using this cascade training technique, which also improves convergence.

b. ESPCN(Efficient Sub-Pixel Convolutional Neural Network): ref add

In order to overcome the drawbacks formed in the traditional super-resolution models, Nay, J. et al. proposed the Efficient Sub-Pixel Convolutional Neural Network (ESPCN)[6] which is an improved architecture used for single-image super-resolution (SISR). ESPCN analyzes the images at their native low resolution for the greater part of its network, in contrast to earlier approaches like SRCNN where it require low-resolution (LR) images to be pre-upscaled using techniques like bicubic interpolation. This architecture is seemed to be more effective and scalable for high-resolution tasks since it uses less computing and memory overhead.

By including a sub-pixel convolution layer in the network's end, the ESPCN model may achieves its efficiency. In order to avoid the necessity for the previous upscaling operations, this layer tries to manages the upscaling operation by learning how to reorganize feature maps into a high-resolution (HR) output. The network can effectively learns the image features that is needed to perform reconstruction while lowering filter size requirements and computing costs by retaining the input low-resolution throughout the first several layers.

The introduction of the Structural Similarity Index Measure (SSIM) as a loss function, instead of the widely using Mean Squared Error (MSE), is seems to be one of the research's major contributions. MSE can reduces pixel-by-pixel variations, but it may ignores perceptual characteristics like texture and structural features, which might be frequently leading to outputs that are fuzzy. However by emphasizing structural similarity and human visual perception SSIM is specifically made to quantify perceived picture quality. The produced results shows sharper edges and more distinct textures which is indicating that this modification enhances the perceived quality of super-resolved images.

By showing how the ESPCN model effectively manages larger image sizes without seeing a linear increase in processing costs, the authors demonstrate the scalability of their method. For the scenarios in which image quality is considered to be more important than numerical accuracy, the incorporation of the SSIM loss function further reinforces ESPCN's advantage in producing visually relevant HR images.

Consequently, by resolving computational inefficiencies and enhancing perceptual quality, ESPCN offers a substantial advancement in the SISR. It differs from the conventional models due to its novel sub-pixel convolution layer and also the incorporation of SSIM as a loss function, which makes it as an appealing option for applications which requires high-quality image reconstruction[24] , such as digital media restoration, medical imaging, and video streaming.

c. **VDSR (Very Deep Super-Resolution):**

In 2016, Very Deep Super-Resolution introduced by Kim et al.[25]. In order to capture more complex image features and enhance the quality of the upscaled images, the fundamental principle of the VDSR is to use the deeper convolutional neural network (CNN) in comparison to previous the model, such as that of SRCNN. The unique features of the VDSR are its depth, residual learning approach, and its application of quick convergence methods, which combine to produce an accurate and an effective system.
Twenty convolutional layers make up the VDSR network architecture, which has a significantly deeper structure than that of the three levels in SRCNN. Bicubic interpolation is used to upsample a low-resolution (LR) image before it is sent into the VDSR. It has 20 convolutional layers and 64 feature maps in each,and the network has a filter of size 3x3 . The Rectified Linear Unit (ReLU) activation function is used in each layer[26]. It tackles the superresolution task with residual learning. By simplifying the intended output, residual learning increases the network's convergence speed and facilitates the training of deeper models. Finally instead of immediately anticipating the high-resolution image, the network produces the residual image, which is the difference between the high-resolution (HR) and low-resolution images. The model is better able to concentrate on picking up the minute details that must be included in the upsampled image attributable to this method.
VDSR is able to reconstruct images with finer details and with greater sharpness by employing a deeper network, particularly in areas with intricate textures. It enhances objective measurements and visual quality by learning finer image structures[27]. Even though VDSR requires more computing power than shallow networks, the total training time is decreased by the residual learning strategy and the application of bicubic interpolation preprocessing. Once trained, the model executes rather quickly.

d. **RCAN(Residual Channel Attention Network):**

In 2018, The Residual Channel Attention Network (RCAN) is a highly advanced architecture introduced by Zhang, Yulun, et al.[8] which is designed for single-image super-resolution (SISR).Using a Residual in Residual (RIR) structure, it makes it possible to build networks with over 400 layers that are incredibly deep. Long skip connections (LSC) and short skip connections (SSC) which are combine to form this structures in numerous residual groups (RGs). The primary architecture that can concentrate on learning the high-frequency features, which are considered to be essential for producing the sharp and the detailed high-resolution outputs due to these connections, which enable the low-frequency inputs bypasses directly through the network.
A Channel Attention (CA) mechanism has been implemented by RCAN to improve the network's learning capacity. Through the modeling of interdependencies between channels, this method adaptively rescales the significance of channel-wise data. The CA mechanism greatly enhances the network's capacity to recover fine textures and features by emphasizing on the most relevant factors which increases its efficiency for hard super-resolution tasks.

To enhance resolution, the RCAN design uses an effective upscaling module in in addition to incorporating shallow and deep feature extraction modules. Recovering features from the low-resolution inputs which is difficult to extract is made possible by using its capacity to ignore the unimportant low-frequency information and concentrate on the crucial high-frequency features. Extensive testing on this typical super-resolution benchmarks has shown that

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

RCAN produces visually appealing results and achieves a state-of-the-art performance in terms of its accuracy (PSNR) and the structural similarity (SSIM).
With its deep architecture, residual connections, and channel attention, RCAN is a potent tool for super-resolution of high-quality images, surpassing many other approaches in terms of both subjective visual quality[28] and objective criteria.

e. **EDSR (Enhanced Deep Super-Resolution Network):**

A major development in the field of superresolution is the Enhanced Deep Super-Resolution Network (EDSR), which was unveiled by Lim et al. in 2017[29]. By eliminating extraneous elements like batch normalization the EDSR improves upon the performance and efficiency of previous models like that of SRCNN and VDSR. When it comes to precision and image quality EDSR has raised the bar in single-image superresolution (SISR) by producing cutting-edge outcomes, particularly at larger upscaling factors.
The fundamental building unit of EDSR is made up of several residual blocks. Two convolutional layers and the ReLU activation function are present in every residual block. EDSR does not include batch normalizing layers, in contrast to previous models such as VDSR. By decreasing the network's ability to learn the intricate mappings required for superresolution, batch normalization reduces network performance. Sharper results and improved generalization are also achieved by removing batch normalization.
In contrast to VDSR, EDSR introduces more filters or feature maps into each of the convolutional layers. This network's broader architecture enables it to extract finer-grained and more detailed characteristics particularly in difficult areas like that of edges and textures. Although the depth of EDSR varies, it has 32 residual blocks when it was first implemented. Furthermore, it employs global residual learning, which makes it possible for the model to learn both low- and high-frequency components more successfully and aids in stabilizing the network's training. The network's architecture allows it to handle various scaling factors such as 2x, 4x, and 8x, without requiring the model to be retrained for every upscaling operations[30][31]. EDSR is more adaptable than that of earlier models since it has its capacity for generalization.
More memory and processing power are needed for the EDSR's deeper and wider design, particularly during its training. The implementation of L1 loss, a pixel-wise loss, can still lead to some oversmoothing even though EDSR obtains high PSNR and SSIM scores, especially in areas with complex textures.

f. **SAN (Second-order Attention Network)**

An innovative single-image super-resolution (SISR) architecture called the Second-Order Attention Network (SAN)[9] aims to improve feature representation and correlation learning in convolutional neural networks (CNNs). The performance of traditional CNN-based SISR models is limited because they frequently provide emphasis to deeper or wider network designs while ignoring the innate correlations among intermediate elements. The SAN framework presents sophisticated strategies to enhance the representational power and prejudiced capability of SISR networks with the aim to tackle such problems.

The Second-Order Channel Attention (SOCA) module, a fundamental part of SAN, leverages second-order feature statistics to adaptively resize channel-wise features. In contrast to earlier techniques that depend on first-order statistics, like global average pooling, SOCA optimizes the ability of the network to concentrate on relevant characteristics by capturing more intricate feature interdependencies. A clearer and more realistic high-resolution image is generated as a result of improved high-frequency texture recovery.

It also enhances SOCA with a Non-Locally Enhanced Residual Group (NLRG) structure. By integrating non-local operations, NLRG enables the network to simulate dependencies outside of local receptive fields by capturing long-distance spatial contextual information. Furthermore, by avoiding low-frequency input, share-source residual groups (SSRG) with skip connections enable better reliable and efficient training of very deep networks.

Shallow feature extraction, deep feature extraction employing NLRG, upscaling and reconstruction are the processes which compose up the SAN architecture. Together, these components enhance pertinent traits, rebuild high-resolution outputs, and retrieve rich information from low-resolution inputs. Additionally, an iterative matrix normalization technique has been implemented into the architecture for effective correlation computing, which speeds up GPU training.

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

In conclusion, by incorporating the non-local operations and the second-order attention mechanisms, the SAN framework provides a noteworthy breakthrough in SISR. SAN sets a new standard for super-resolution challenges by focusing on spatial context and feature interdependence, which yields better results while effectively balancing model complexity and performance.

g. **HAN(Holistic Attention Network)**

Niu, B. et al. [10] designed an innovative deep learning framework known as the Holistic Attention Network (HAN) to overcome the drawbacks of single-image super-resolution (SISR) by implementing attention mechanisms that adaptively highlight important features throughout channels, spatial dimensions, and hierarchical levels.

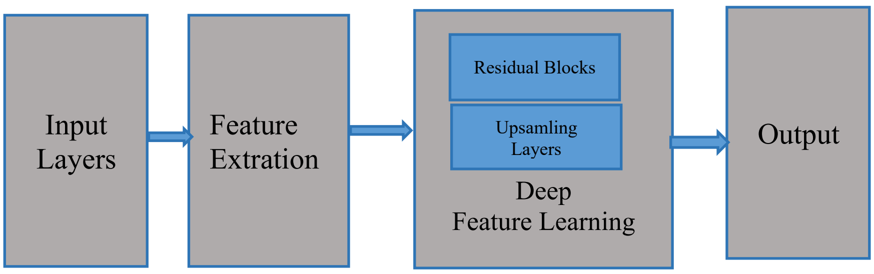

Fig 3.3: Basic architecture of HAN

The Layer Attention Module (LAM) and the Channel-Spatial Attention Module (CSAM) are two novel modules that comprises the foundation of HAN. The residual groups in the network generate hierarchical features whereas the LAM records the associations between these features. It also allows the model to suppress duplicate layers and emphasize informative ones by assigning weights to features according to their significance. By utilizing long-distance dependencies across layers—a feature that typical dense and skip connections ignore—this successfully increases the model's representational strength.

The CSAM improves feature discrimination by simulating spatial locations and interdependencies both inside and across channels. CSAM integrates both spatial dimensions and channels, in contrast to traditional attention mechanisms that only concentrate on one or the other. This guarantees that the network captures contextual linkages and fine-grained features in each feature map. This module is applied to the deepest layer's output, allowing the model to balance improved performance with computational expenses.

HAN's architecture seamlessly integrates these attention techniques with an RCAN-inspired core network. It applies holistic attention using LAM and CSAM, retrieves hierarchical features using residual groups, and then uses a sub-pixel convolution layer for upscaling to rebuild the HR image. A more reliable and effective model is produced by long skip connections, which also stabilize training and maintain low-frequency information.

Therefore, by including holistic attention mechanisms that adaptively improve feature acquisition, the HAN offers a significant development in the SISR. The model may capture important dependencies at several levels through the combination of LAM and CSAM, which in a way improves accuracy and visual quality. HAN is considered to be a potent tool for applications like digital content improvement, video surveillance, and medical imaging[32] that demand high-quality picture reconstruction because of these advancements.

h. **CARN(Cascading Residual Network)**

Namhyuk et.al[11] proposed the lightweight CARN-M version and the Cascading Residual Network (CARN) for single-image super-resolution (SISR) . These models tackle the problem of attaining high accuracy while preserving computational efficiency, which qualifies them for practical uses such as streaming video and mobile devices. CARN enhances feature representation by utilizing cascading connections on both a local and global level, which makes it possible to rebuild the fine details in images and can deal with the problems like vanishing gradients.

Multi-level feature representation can be made possible by CARN, which introduces cascading connections at both the local and global levels. Intermediate outputs within the block can be connected by local cascading, whereas the outputs across blocks are connected by global cascading. This architecture helps to ensure the enhanced feature fusion, smoother gradient flow during training, and more successful image detail reconstruction. For accurate image restoration and effective training, the network utilizes residual learning.

With the help of effective residual blocks within group and pointwise convolutions and recursive weight-sharing, the light-wight CARN further optimizes the design, drastically lowering the number of parameters and operations without compromising the performance. By using a late upsampling technique and multi-scale learning, the models enable a single model to handle many upscaling factors.

Numerous analyses on benchmark datasets show that CARN and Light weight CARN (CARN-M) perform better than or are on par with the most advanced models in terms of computing efficiency and accuracy . Due to scenario the design work especially well in applications where performance, speed, and power consumption need to be matched.

The usefulness of the models is especially evident in mobile devices, where energy and computational efficiency are crucial. These models can upscale compressed videos to high quality with minimal delays. Deployment is made further simpler by the multi-scale capabilities[33][34], which does away with different models.

i. RDN(Residual Dense Network)

In order to overcome the difficulty of completely utilizing hierarchical features for high-quality image reconstruction, Zhang et al.[12] developed the Residual Dense Network (RDN), a very successful design for single-image super-resolution (SISR). Conventional SISR techniques based on deep learning, such those that use dense or residual blocks, are unable to optimize feature use across all convolutional layers. By incorporating Residual Dense Blocks (RDBs), which extract, improve, and adaptively retain hierarchical features by combining dense connectivity, local feature fusion, and local residual learning makes the RDN gets over this restriction.

The fundamental components of RDN are Residual Dense Blocks (RDBs). They are made up of different layers that are densely connected, with each layer's output connecting to every layer that comes after it in the block. Thus complete local feature extraction can be ensured by this high connection. Every RDB also uses the Local Residual Learning (LRL), which may stabilizes training and improves gradient flow by introducing a residual link across the block. Additionally, to improve feature representation and network capacity, the architecture incorporates Local Feature Fusion (LFF), which combines features from previous and current layers in an adaptive manner.

RDBs use a contiguous memory mechanism that enables each layer in the following RDB to directly access the output state of the previous RDB. Learning and performance can be enhanced by this method which also guarantees effective gradient propagation and information flow throughout the network.

RDN utilizes Global Feature Fusion (GFF) to merge local features that have been extracted using RDBs. By combining hierarchical characteristics from every RDB in the network this procedure allows the model to use both high-level and low-level data to accurately reconstruct images. Global Residual Learning (GRL), which combines the deep features from the RDBs with shallow features taken from the input to develop global dense features helps to improves GFF even further.

RDN is appropriate for real-world uses such as medical imaging, video improvement, and real-time super-resolution in mobile devices due to its effective architecture[35]. By utilizing this feature fusion and dense connectivity, the network may effectively expand with depth and width, enabling it to adjust to evolving computational and performance demands.

j. DRCN (Deep Recursive Convolutional Network)

For single-image super-resolution (SISR), Kim et al. [13] propose the Deeply-Recursive Convolutional Network (DRCN), which tries to emphasizes the deep recursion to boost field responsiveness while preserving the efficiency. A broad receptive field may be achieved by DRCN by applying the single convolutional layer recursively up to 16 times. This allows for the improved recovery of the high-frequency information without need for extra parameters.

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

Recursive-supervision, which helps to oversees all intermediate recursive outputs, and skip-connections, which may connect the input straight to the reconstruction layer, are considered to be significant innovations. By handling training the challenges including vanishing and bursting gradients, these methods helps to provide stable optimization and improved feature learning.

An embedding network for feature representation extraction, the recursive inference network for contextual information analysis, and the reconstruction network for high-resolution output generation constitute the DRCN architecture. Together, these elements facilitate effective processing for tremendous image contexts[36].

The design helps to surpasses current approaches in terms of accuracy and visual quality and shows superior results across standard benchmarks. In addition to being resource-efficient, the suggested framework uses fewer parameters than that of the existing deep networks, which makes it ideal for picture restoration tasks like artifact removal and denoising.

k. **DRRN (Deep Recursive Residual Network)**

Ying Tai et al. [14] presented a unique method for single-image super-resolution (SISR) called the Deep Recursive Residual Network (DRRN). DRRN tackles the problem of building deep networks with low memory and computational needs while maintaining good performance. By using recursive and residual learning, DRRN achieves better outcomes with fewer parameters than previous approaches that need large numbers of parameters for greater depth.

Global residual learning (GRL) and local residual learning (LRL) are frameworks utilized in the DRRN architecture. The LRL uses recursive blocks to enhance gradient flow and also to preserve high-frequency details, whereas GRL makes training easier by learning the difference between the input low-resolution (LR) and the target high-resolution (HR) image output . DRRN may reach depths of up to 52 layers without substantially increasing the number of parameters through recursive learning, which improves the model even more by reusing parameters across numerous residual units[37][38].

The multi-path residual block structure, which allows several residual units to share an input within each block, is vital for the design of DRRN. This framework utilize gradient backpropagation which is more efficient and allows the model to learn more intricate features. The multi-path strategy works effectively especially well for small models given that it is less likely to overfit than that of single-path techniques.

Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) are seems to be higher with DRRN than with other approaches, indicating the state-of-the-art performance across multiple datasets. Additionally, compared to similar deep learning techniques like VDSR and DRCN, DRRN uses up to 14 times less parameters to attain these outcomes.

The design helps to improves the versatility and scalability of DRRN. The model might be utilized for various application by modifying it without compromising the speed by varying the number of recursive blocks and units. Numerous tests help to confirm the effectiveness of DRRN, and qualitative evaluations reveal crisp and appealing reconstructions.

Therefore, DRRN considered as a very successful and efficient option for SISR. It sets a new standard in the sector with its deep yet lightweight architecture, recursive and residual learning techniques, and exceptional performance metrics.

l. **SRGAN (Super-Resolution Generative Adversarial Network)**

One of the most innovative models in single-image super-resolution (SISR) was presented by Ledig et al. in 2017[15] the Super-Resolution Generative Adversarial Network (SRGAN). To improve perceptual quality, SRGAN uses a generative adversarial network (GAN), in contrast to typical deep learning techniques that only aim to minimize pixel-wise discrepancies between the super-resolved image and the high-resolution ground truth[39][40]. SRGAN concentrates on creating aesthetically pleasing, high-frequency details, which can lead to clearer and more appealing pictures[41]. This is in contrast to previous models like SRCNN, VDSR, and EDSR, which place more emphasis on quantitative indicators like Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM).

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

SRGAN has been invented to overcome this limitation by introducing a perceptual loss function that, when coupled with an adversarial loss, results in images that are finer and more realistic. By integrating a discriminator network to the generator, SRGAN produces images that match the ground truth in pixel space and reproduce the perceptual features of real high-resolution images.

SRGAN is made up of two fundamental components, namely the discriminator and the generator. Together, they form an adversarial network where the generator attempts to produce high-quality images from low-quality images while the discriminator strives to distinguish between generated super-resolved images and real high-resolution photos.

In addition to removing pixel-wise differences, SRGAN introduced the idea of perceptual loss, which compares high-level feature maps from a VGG network that has been previously trained. The resulting images are more visually appealing, crisp, and have a greater degree of texture. SRGAN, which uses a GAN framework, is utilized to train the generator for generating images that are exactly like real high-resolution images. The adversarial loss encourages the generator to deliver high-frequency details that are frequently ignored in pixel-based super-resolution models.

SRGAN was developed to get over this restriction by adding a perceptual loss function that, when coupled with an adversarial loss, produces finer, more genuine images. SRGAN creates images that replicate the perceptual characteristics of actual high-resolution images and match the ground truth in pixel space through integrating a discriminator network into the generator.

Training the discriminator and generator at the same time can be challenging since an imbalance among them could lead to problems like model collapse, when the generator generates a small range of images. Proper hyperparameter adjustment is required for stability throughout training. SRGAN training is also computationally expensive.

m. **ESRGAN(Enhanced Super-Resolution Generative Adversarial Network**)

The SRGAN design is greatly improved in various ways by ESRGAN. It improves the generator's architecture and the loss functions while maintaining the discriminator adversarial network configuration which helps to generate more realistic and visual appealing super-resolved images[18]. Furthermore, by adding more features like the Residual-in-Residual Dense Block (RRDB), Relativistic GAN loss, and an enhanced visual loss function, ESRGAN improves detail preservation and gets rid of artifacts[42][43].

When compared to simpler models used in super- resolution like EDSR and the models that use quantitative metrics that emphasize pixel-wise precision, like PSNR or SSIM, might not indicate any gain over ESRGAN. However, ESRGAN produces images that are most probably cleaner and more aesthetically pleasing to eyes than those produced by SRGAN and other algorithms. ESRGAN specializes in generating images which conserves natural textures, sharp edges, and finer details. It seems to be useful especially for image enhancement applications where visual appeal is more crucial since it can preserve high-frequency textures like hair, grass, and fur. The remarkable visual quality of ESRGAN's that produced is seems to be one of its primary strengths.

Unlike SRGAN, ESRGAN enhances convergence and decreases difficulties with training by stabilizing the training procedure with the help of the Relativistic Discriminator and deeper residual blocks.

The two most significant challenges with ESRGAN are believed to be the model's complexity and the training computation time. Despite ESRGAN's ability to generate realistic textures, adversarial training may occasionally induce undesired artifacts. Since these artifacts might not be apparent in the ground truth image, there may be arbitrarily distortions.

4. COMPARISON BETWEEN THE DIFFERENT SUPER-RESOLUTION MODEL

The following table compares various models used in Super-Resolution technique. The comparison focused on its architecture, accuracy, efficiency and its key innovation aspects.

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The SSIM (Structural Similarity Index) and The PSNR (Peak Signal-to-Noise Ratio) can be used to determine the accuracy of various models that can be used in super-resolution. Better fidelity to the ground truth can be shown by the high PSNR values, whereas perceptual similarity may measured in SSIM, where higher values indicate better structural consistency. While higher accuracy values often used to signify the greater reconstruction fidelity, they do not account for perceptual quality, which is crucial for the GAN-based models such as SRGAN and ESRGAN. In terms of perceptual quality, the models give greater prominence to producing aesthetically beautiful images than to optimizing numerical measures such as SSIM and PSNR.

Sl No.	Models	Architecture	Accuracy	Efficiency	Key Innovations
1	**SRCNN**	**3 Layers**	**Low**	**Low and Fast inference.**	**First deep learning-based SR model, and have simple design.**
2	**ESPCN**	**Shallow**	**Low to Moderate**	**High and memory cost.**	**Sub-pixel convolution for efficient upscaling, and it avoids early dimensional expansion.**
3	**VDSR**	**20 Layers**	**Moderate**	**Moderate but larger memory usage.**	**Residual learning for deeper networks and improved gradient flow.**
4	**RCAN**	**Very Deep (About 400+ layers)**	**High**	**Low and Heavy Computation.**	**Residual groups with channel attention mechanisms that has been used for fine-grained features enhancement.**
5	**EDSR**	**Deep (About 16+**	**High**	**Moderate but resours**	**Optimized deep network by**

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

		layers)		**e intensive**	**removing Batch Normalization (BN) can be used for better image quality.**
6	**SAN**	**Deep (About 30+ layers)**	**High**	**Moderate but resourse intensive**	**Second-order attention mechanisms used for better channel correlation and feature refinement.**
7	**HAN**	**Very Deep (About 400+ layers)**	**High**	**Moderate**	**Layer attention used for hierarchical features and channel-spatial attention may used for local/global feature focus.**
8	**CARN**	**Shallow (About 10+ layers)**	**Moderate**	**High and Mobile friendly**	**Lightweight cascading residual network may used for efficiency in resource-constrained environments.**
9	**RDN**	**Deep (About 16+ layers)**	**High**	**Moderate**	**Residual dense connections for effective feature reuse and high-quality reconstruction.**
10	**DRCN**	**Deep recursi**	**Moderate**	**High parame**	**Recursive convolutiona**

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

		ve (about 16 rec. layers)		**ter effcienc y**	**l layers with skip connections can be used for large receptive fields and efficient training.**
1 1	**DRRN**	**Deep recursi ve (about 52 rec. Layers)**	**Mod erate**	**High parame ter effcienc y**	**Combines recursive and residual learning can be used for better convergence and high-quality results.**
1 2	**SRGAN**	**Deep (Consis t of Genera tor and Discri minato r)**	**Low for PSN R/ SSIM , but high perc eptu al quali ty.**	**Modera te and GAN require s tuning.**	**First GAN-based SR model, focused on perceptual realism rather than numerical accuracy.**
1 3	**ESRGA N**	**Deep (Consis t of Genera tor and Discri minato r)**	**Low for PSN R/ SSIM , but high perc eptu al quali ty.**	**Modera te and improv ed efficien cy.**	**Enhanced version of SRGAN which has Residual-in-Residual Dense Blocks (RRDB) for realistic texture generation.**

Table 4.1: Comparison between various SR-Models

Comparison Based on PSNR and SSIM

The following table gives a comparison of super-resolution models at a scaling factor of ×4(×4 means the LR images are scaled down by a factor of 4 and then reconstructed and compare the same with original) on the standard benchmark datasets such as Set5, Set14, Urban100, B100 based on PSNR and SSIM.

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

Set 5 is a data set made up of 5 images that feature a variety of objects and landscapes. This dataset is seems to be more clear-cut, uncomplicated, and has distinct textures, which makes to helpful for observing fundamental SR performance. In comparison to Set5, Set14 offers a larger test set of 14 images, which aids for evaluating generalization to a variety of textures and patterns. The purpose of Urban100 is said to be test SR techniques on repeating textures and structural patterns like that of grids and lines, that are present in urban environments. It consists of 100 photos of urban sceneries. The Berkeley Segmentation Dataset, or B100, is another collection of 100 natural scene photos that may be used to test SR models on a variety of natural situations with different textures.

Sl No.	Models	Set5 (PSNR/ SSIM)	Set14 (PSNR/ SSIM)	Urban100 (PSNR/ SSIM)	B100 (PSNR/ SSIM)
1	SRCNN	30.48/ 0.862	27.49/ 0.751	24.52/ 0.722	26.90/ O.710
2	ESPCN	30.90/ 0.870	27.70/ 0.760	24.80/ 0.730	27.00/ 0.720
3	VDSR	31.35/ 0.883	28.02/ 0.789	25.18/ 0.752	27.29/ 0.725
4	RCAN	32.63/ 0.900	28.87/ 0.789	26.82/ 0.808	27.77/ 0.743
5	EDSR	32.46/ 0.898	28.80/ 0.788	26.64/ 0.803	27.71/ 0.742
6	SAN	32.64/ 0.900	28.92/ 0.790	26.79/ 0.806	27.78/ 0.744
7	HAN	32.64/ 0.901	28.90/ 0.790	26.85/ 0.807	27.80/ 0.744
8	CARN	31.92/ 0.890	28.42/ 0.774	25.62/ 0.764	27.44/ 0.735
9	RDN	32.47/ 0.899	28.81/ 0.787	26.61/ 0.802	27.72/ 0.743
10	DRCN	31.53/ 0.884	28.02/ 0767	25.14/ 0.762	27.23/ 0.728
11	DRRN	31.68/ 0.888	28.21/ 0.772	25.44/ 0.762	27.38/ 0.728
12	SRGAN	Perceptual Quality	Perceptual Quality	Perceptual Quality	Perceptual Quality
13	ESRGAN	Perceptual Quality	Perceptual Quality	Perceptual Quality	Perceptual Quality

5. RESULTS AND DISCUSSION

Significant distinctions in performance among metrics and applications were found when analyzing various Super-Resolution models which includes SRCNN, ESPCN, VDSR, RCAN, EDSR, SAN, HAN, CARN, RDN, DRCN, DRRN, SRGAN, and ESRGAN. As a pioneering deep learning model the SRCNN performed mediocrely but had trouble capturing high-frequency features. In a similar vein, ESPCN, which prioritizes sub-pixel convolutions for effective upscaling, demonstrated faster inference time but less precise outputs when compared to more recent models. VDSR performed superior to these previous models by increasing structural accuracy and producing finer outcomes by utilizing a deeper architecture.

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

Advanced architectures like RCAN, EDSR and RDN shows better performance in terms of SSIM and PSNR with EDSR attaining high accuracy through deeper but simpler network topology designs. Medical imaging and other applications needing complicated reconstructions can benefit from RCAN's ability to handle tiny details and complex textures thanks to its attention mechanism. Similarly, RDN generated detailed reconstructions by effectively capturing hierarchical features, especially in high-resolution applications like satellite images.

For real-time applications, CARN was the best lightweight model because it struck a compromise between efficiency and performance. DRCN and DRRN, on the other hand, used recursive layers to lower model parameters while retaining competitive accuracy however, their performance was inferior to that of more recent models such as RCAN and EDSR.

Instead of focusing on precise pixel accuracy, GAN-based models like SRGAN and ESRGAN prioritized perceptual quality. The most aesthetically accurate results were produced by ESRGAN, an improvement on SRGAN, which produced vibrant textures and high SSIM ratings. However these models were computationally intensive and sometimes introduced artifacts, particularly in areas that were uniform or repeating. By combining sophisticated attention mechanisms, SAN and HAN went beyond more straightforward GAN-based methods in maintaining both global and local consistency, challenging the limits of perceptual realism and detail recovery.

Qualitative comparisons indicated that less complex models, such as SRCNN and ESPCN, generated outputs that were smooth but less detailed, making them appropriate for situations requiring limited resources. In contrast the RCAN, EDSR, and ESRGAN stood out for their ability to rebuild delicate textures and sharp edges. These results highlight the trade-off between output quality and computational expense. For embedded systems the lightweight models like CARN are ideal whereas high-performance models like RCAN and ESRGAN are more appropriate for offline computing in high-end applications.

The study highlights the significance of choosing an SR model according to their particular application and available resources. To bridge the gap between the efficiency and the quality future research could concentrate on merging hybrid loss functions with lightweight attention mechanisms, making advanced SR approaches more accessible for a variety of use applications.

6. APPLICATIONS, FUTURE DIRECTIONS AND EMERGING TRENDS

Super-Resolution (SR) may help to improve the visual clarity and image details of the low resolution images which makes the models more useful in numerous applications. In medical imaging the SR can enhances the resolution of the medical scans like MRIs, CTs, and X-rays, facilitating precise diagnosis and early detection of disease. In satellite and remote sensing, SR may improves the quality of satellite images which helps in urban planning, disaster relief and environmental monitoring. Low-resolution video footage can be up-scaled for security and surveillance purposes to facilitate identification and analysis. SR is also used extensively in the entertainment sector for video up-scaling, which converts old films and standard-definition material into high-definition forms. Furthermore, SR is used in digital art restoration, where deteriorated or old artwork is recreated with improved visuals, and in gaming, where textures are up-scaled in real-time to elevate visual experiences.

Significant advancement is expected in SR in the future, especially in balancing computational efficiency and quality. The prevalence of lightweight SR models is expected to be increase since they allow real-time applications on the edge devices which includes smartphones and embedded systems, without sacrificing output quality. Additionally, unsupervised and self-supervised learning techniques will solve the shortage of paired high-resolution and low-resolution datasets, expanding the use of SR in domains involving historical document restoration and astronomy.

The confluence of SR with multi-modal learning and domain-specific adaptation where SR models are customized for particular applications, such driver less cars or medical imaging are emerging developments. Blockchain technology with federated learning will provide the safe data exchange and privacy in various applications that may handle sensitive information. Moreover, by upscaling the real-time graphics, SR may be crucial in improving the virtual reality (VR) and the augmented reality (AR) experiences. There are still issues to be resolved, such as lowering artifacts in GAN-based models, dealing with the high computational cost of sophisticated structures and guaranteeing that SR is generalizable across a variety of image in various domain. Future studies will probably concentrate on resolving the above issues so that SR may be used more widely and with greater adaptability.

7. CONCLUSION

The variety of capabilities and trade-offs present in these models are demonstrated by the comparison of several Super-Resolution (SR) approaches. While simpler models like SRCNN and ESPCN offer faster and more computationally efficient solutions, these are provided at the expense of output quality and are appropriate for real-time or low-resource applications. Advanced designs like RCAN, EDSR and ESRGAN are perfect for applications which needs high-fidelity outputs like media restoration, medical imaging, and satellite mapping as they shows notable advances in reconstructing fine details and perceptual quality.

Although they require more processing power, GAN-based models—particularly ESRGAN—are excellent at producing textures that are visually realistic. These models work best in offline processing applications like digital content production or film restoration where quality is crucial. Recursive methods like DRRN and lightweight models like CARN provide a balance between efficiency and performance, which makes them useful for real-time augmentation jobs or embedded systems.

The study emphasizes on the significance of choosing SR models in accordance with the particular application requirements such as demands for quality, real-time limitations, and computational resources. The development of the hybrid models that combine the effectiveness of lightweight structures with the caliber of sophisticated attention mechanisms may be the main emphasis of future SR developments which would increase the accessibility and industry-wide relevance of SR techniques.

REFERENCES:

[1] Chen, H., He, X., Qing, L., Wu, Y., Ren, C., Sheriff, R. E., & Zhu, C. (2022). Real-world single image super-resolution: A brief review. Information Fusion, 79, 124-145.

[2] Z. Wang, J. Chen and S. C. H. Hoi, "Deep Learning for Image Super-Resolution: A Survey," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 10, pp. 3365-3387, 1 Oct. 2021, doi: 10.1109/TPAMI.2020.2982166. keywords: {Deep learning;Degradation;Animals;Benchmark testing;Measurement;Image super-resolution;deep learning;convolutional neural networks (CNN);Generative adversarial nets (GAN)},

[3] Anwar, S., Khan, S., & Barnes, N. (2020). A deep journey into super-resolution: A survey. ACM Computing Surveys (CSUR), 53(3), 1-34.

[4] Farsiu, S., Robinson, D., Elad, M., & Milanfar, P. (2004). Advances and challenges in super-resolution. International Journal of Imaging Systems and Technology, 14(2), 47-57.

[5] Kumar, Sujith. "V.: Perceptual image super-resolution using deep learning and super-resolution convolution neural networks (SRCNN)." Intell. Syst. Comput. Technol 37.3 (2020).

[6] Nay, J. (2021). Single image super resolution using espcn–with ssim loss. Available at SSRN 3898913

[7] Hitawala, Saifuddin, et al. "Image super-resolution using VDSR-ResNeXt and SRCGAN." arXiv preprint arXiv:1810.05731 (2018).

[8] Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., & Fu, Y. (2018). Image super-resolution using very deep residual channel attention networks. In Proceedings of the European conference on computer vision (ECCV) (pp. 286-301).

[9] Dai, T., Cai, J., Zhang, Y., Xia, S. T., & Zhang, L. (2019). Second-order attention network for single image super-resolution. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 11065-11074).

[10] Niu, B., Wen, W., Ren, W., Zhang, X., Yang, L., Wang, S., ... & Shen, H. (2020). Single image super-resolution via a holistic attention network. In Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XII 16 (pp. 191-207). Springer International Publishing.

[11] Ahn, N., Kang, B., & Sohn, K. A. (2018). Fast, accurate, and lightweight super-resolution with cascading residual network. In Proceedings of the European conference on computer vision (ECCV) (pp. 252-268).

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

[12] Zhang, Y., Tian, Y., Kong, Y., Zhong, B., & Fu, Y. (2018). Residual dense network for image super-resolution. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2472-2481).
[13] Kim, J., Lee, J. K., & Lee, K. M. (2016). Deeply-recursive convolutional network for image super-resolution. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1637-1645).
[14] Tai, Y., Yang, J., & Liu, X. (2017). Image super-resolution via deep recursive residual network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3147-3155).
[15] Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
[16] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of The 32nd International Conference on Machine Learning (ICML),
pages 448–456, 2015
[17] K. He, X. Zhang, S. Ren, and J. Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In IEEE International Conference on Computer Vision (ICCV), pages 1026–1034, 2015.
[18] Wang, Xintao, et al. "Esrgan: Enhanced super-resolution generative adversarial networks." Proceedings of the European conference on computer vision (ECCV) workshops. 2018.
[19] Dong, Chao, et al. "Learning a deep convolutional network for image super-resolution." *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part IV 13*. Springer International Publishing, 2014.
[20] Lv, Y., & Ma, H. (2021, April). Improved SRCNN for super-resolution reconstruction of retinal images. In 2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP) (pp. 595-598). IEEE.
[21] Ooi, Y. K., & Ibrahim, H. (2021). Deep learning algorithms for single image super-resolution: a systematic review. Electronics, 10(7), 867.
[22]Shen, K., Guo, J., Tan, X., Tang, S., Wang, R., & Bian, J. (2023). A study on relu and softmax in transformer. arXiv preprint arXiv:2302.06461.
[23] Ren, Haoyu, Mostafa El-Khamy, and Jungwon Lee. "CT-SRCNN: cascade trained and trimmed deep convolutional neural networks for image super resolution." *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2018.
[24] Xu, M., Du, X., & Wang, D. (2020, March). Super-resolution restoration of single vehicle image based on ESPCN-VISR model. In IOP Conference Series: Materials Science and Engineering (Vol. 790, No. 1, p. 012107). IOP Publishing.
[25] Kim, Jiwon, Jung Kwon Lee, and Kyoung Mu Lee. "Accurate image super-resolution using very deep convolutional networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016
[26] Hitawala, S., Li, Y., Wang, X., & Yang, D. (2018). Image super-resolution using VDSR-ResNeXt and SRCGAN. arXiv preprint arXiv:1810.05731.
[27] Vint, D., Di Caterina, G., Soraghan, J. J., Lamb, R. A., & Humphreys, D. (2019, May). Evaluation of performance of VDSR super resolution on real and synthetic images. In 2019 Sensor Signal Processing for Defence Conference (SSPD) (pp. 1-5). IEEE.
[28] Liu, Y. (2023, October). RCAN based MRI super-resolution with applications. In Proceedings of the 2023 7th International Conference on Electronic Information Technology and Computer Engineering (pp. 357-361).
[29] Lim, Bee, et al. "Enhanced deep residual networks for single image super-resolution." *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2017.
[30] Lin, K. (2022, September). The performance of single-image super-resolution algorithm: EDSR. In 2022 IEEE 5th international conference on information systems and computer aided education (ICISCAE) (pp. 964-968). IEEE.
[31] Kuriakose, B. M. (2023). EDSR: Empowering super-resolution algorithms with high-quality DIV2K images. Intelligent Decision Technologies, 17(4), 1249-1263.

[32] Mehta, S., Lu, X., Weaver, D., Elmore, J. G., Hajishirzi, H., & Shapiro, L. (2020). Hatnet: an end-to-end holistic attention network for diagnosis of breast biopsy images. arXiv preprint arXiv:2007.13007.
[33] Lan, R., Sun, L., Liu, Z., Lu, H., Su, Z., Pang, C., & Luo, X. (2020). Cascading and enhanced residual networks for accurate single-image super-resolution. IEEE transactions on cybernetics, 51(1), 115-125.
[34] Yang, H., Dou, Q., Liu, K., Liu, Z., Francese, R., & Yang, X. (2021). Efficient local cascading residual network for real-time single image super-resolution. Journal of Real-Time Image Processing, 18, 1235-1246.
[35] Kong, D., Gu, L., Li, X., & Gao, F. (2024). Multi-Scale Residual Dense Network for the Super-Resolution of Remote Sensing Images. IEEE Transactions on Geoscience and Remote Sensing.
[36] Tai, Y., Yang, J., & Liu, X. (2017). Image super-resolution via deep recursive residual network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3147-3155).
[37] Panaetov, A., Daou, K. E., Samenko, I., Tetin, E., & Ivanov, I. (2022). Rdrn: Recursively defined residual network for image super-resolution. In Proceedings of the Asian Conference on Computer Vision (pp. 4110-4125).
[38] Yang, C., & Lu, G. (2020). Deeply recursive low-and high-frequency fusing networks for single image super-resolution. Sensors, 20(24), 7268.
[39] Huang, J. H., Wang, H. K., & Liao, Z. W. (2020, October). HFD-SRGAN: super-resolution generative adversarial network with high-frequency discriminator. In 2020 IEEE international conference on systems, man, and cybernetics (SMC) (pp. 3148-3153). IEEE.
[40] Ye, Z., Lyu, F., Li, L., Fu, Q., Ren, J., & Hu, F. (2019, July). SR-GAN: Semantic rectifying generative adversarial network for zero-shot learning. In 2019 IEEE international conference on multimedia and expo (ICME) (pp. 85-90). IEEE.
[41] Zhu, J., Yang, G., & Lio, P. (2019, April). How can we make GAN perform better in single medical image super-resolution? A lesion focused multi-scale approach. In 2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019) (pp. 1669-1673). IEEE.
[42] Bell-Kligler, S., Shocher, A., & Irani, M. (2019). Blind super-resolution kernel estimation using an internal-gan. Advances in Neural Information Processing Systems, 32.
[43] You, C., Li, G., Zhang, Y., Zhang, X., Shan, H., Li, M., ... & Wang, G. (2019). CT super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-CIRCLE). IEEE transactions on medical imaging, 39(1), 188-203.