Journal of Computer Science

Original Research Paper

Image Super-Resolution using Auto-Encoders with Parallel Skip-Connections

Lisha Pulickal Purushothaman and Jayasree Vadakke Kadangote

Department of Electronics and Communication, Government Model Engineering College, Thrikakkara, India

Article history
Received: 19-05-2022
Revised: 25-08-2022
Accepted: 14-09-2022

Corresponding Author:
Lisha Pulickal Purushothaman
Department of Electronics and Communication, Government Model Engineering College, Thrikakkara, India
Email: lishapp@gmail.com

Abstract: Super-resolution images are highly desired when employed for numerous analytical purposes and obviously because of their superior attractive visual effect. To create a High-Resolution (HR) image from one or more Low-Resolution (LR) images, an image super-resolution technique is used. When dealing with natural environments and settings, it is not always easy to access high-resolution images. The main barriers to the same are limitations in acquisition methods. In the domains of forensics investigation, remote sensing, digital monitoring, and medical imaging high-resolution images are always required. Modern methods using deep learning models have improved performance when compared to classic image processing methods. Using a convolution auto-encoder architecture with six parallel skip connections, a novel technique for improving lower solution natural images of size 256 × 256 to high-resolution images is proposed in this paper. The use of parallel skip connections between the encoder and decoder allows the network to reconstruct high-resolution images while still extracting relevant features from low-resolution images. Furthermore, the network is skillfully tuned using the right filters and regularization techniques. To create high-resolution output images, the decoder component makes use of the features' reduced representation in latent space. The model was evaluated using DIV 2K, CARS DATA, Set5, Set14, and the General data set after being trained using various data sets like CARS DATA and DIV 2K. Peak Signal to Noise Ratio, Structural Similarity Index, Mean Squared Error, and model behavior on numerous data sets are used to compare the proposed method to various current methods. Results reveal that the suggested model works better than the existing method and is reliable.

Keywords: Image Super-Resolution, Convolution Auto-Encoder, Feature Extraction, Encoder, Latent Representation, Decoder

Introduction

A good quality, clear image is required for viewing and analyzing images for a variety of purposes, including satellite imaging, medical imaging, improving face and text images, and enhancing compressed images. In both spatial and frequency domains, there exist various approaches for image enhancement. In the spatial domain, enhancement techniques work directly with pixels; however, in the frequency domain, they work through transform coefficients. The smallest detectable detail in a visual presentation is characterized as image resolution. Resolution is divided into four categories in digital image processing: Spatial resolution, brightness resolution, temporal resolution, and spectral resolution. Using one or more low-resolution images, the Super-Resolution (SR) approach creates an image with a high spatial resolution. There are two different kinds of super-resolution techniques such as single frame super-resolution and multi-frame super-resolution. Both traditional and deep learning techniques have been used to build the SR models in both of these categories. For each super-resolution category in the traditional technique and raw Gilman *et al*. (2008) noticed various strategies using interpolation, learning techniques, and reconstruction techniques. The most popular interpolation methods are the nearest neighbor, linear and cubic splines. The weighted average of adjacent LR pixels was used by these approaches to predict unknown HR pixels. While

interpolation-based techniques are rapid and simple, Ledig *et al*. (2017) assert that it is more challenging to evaluate the image's precision, leading to a blurring of promising features and high-frequency components in a sample image. Algorithms based on the learning technique map the LR and HR images using an external data set. Nasrollahi and Moeslund (2014) observed that the details of the HR image are obtained utilizing some prior knowledge in reconstruction methods. Using conventional techniques, images cannot be enlarged without sacrificing image quality.

Super-resolution models based on deep learning have recently outperformed traditional methods in all applications requiring image analysis. This study proposed a convolution auto-encoder-based SR model with parallel skip connections for extracting all prominent features of the input image and coding into a compressed latent representation which is used to reconstruct a good quality image using a decoder network.

Figure 1 summarizes super-resolution techniques based on deep learning algorithms. They are categorized into pre-up sampling, post-up sampling, residual recursive architecture, progressive reconstruction networks, multi-branch networks, attention networks, and generative adversarial network-based models.

Pre-Up Sampling Approach

Convolution neural network-based Super-Resolution (SRCNN) and very deep super-resolution are well-known SR models in this category. In these models, classical SR algorithms are integrated with deep learning methods for reconstructing good-quality images. SRCNN is the first convolution neural network-based superre solution model introduced by Dong *et al*. (2015). Convolution layers are utilized in this model to achieve patch extraction, nonlinear mapping, and reconstruction. However, this model struggles to transmit the input image's texture and structure to deeper layers. A very deep super-resolution model proposed by Kim *et al*. (2016), employed a deeper network (20 layers) with a suitable number of filters and lower size kernel. Since the network is deeper, this model takes significantly longer training time and consume more memory, and also reconstructed image suffers from artifacts or blur.

Post-Up Sampling Approach

A low-dimensional space is used for feature extraction in the post-up-sampling approach and up-sampling is performed at the end of the procedure. The Efficient Subpixel Convolutional Network (ESPCN) and the Fast Superre Solution Convolutional Neural Network (FSRCNN) are two well-known super-resolution models in this area. The deconvolution procedure is used in Dong *et al*. (2016) suggested FSRCNN and their model's up sampling. In their model, Shi *et al*. (2016) introduced ESPCN, which was used for up-sampling with sub-pixel convolution operation. This technique avoids the memory and time limitations of the prior model because the up-sampling operation is conducted toward the end of the process, but the regenerated image gets blurred for high magnification orders.

Progressive-Up Sampling

In this approach, many phases of up-sampling and feature extraction are performed for reconstructing a good-quality image. Lai *et al*. (2017) proposed a Laplacian pyramid Super-Resolution Network (Lap SRN) which employed feature embedding and feature up sampling in the feature extraction stage and the reconstruction stage. Before up-sampling, in this model, features are extracted and they are then mixed with the interpolated image to produce the HR image. The features extracted are utilized again to refine the high-resolution image that will be used in the next stage. Since this model used many stages of feature extraction and reconstruction, the network becomes complex and training consumes more time.

Residual Network

Enhanced Deep residual SR(EDSR) and Multi-scale SR (MDSR) are residual network-based SR models. Lim *et al*. (2017) introduced EDSR and MDSR which implemented both single-scale and multi-scale superre solutions through multiple residual blocks in a single model and eliminated the batch normalization layer. Later this model was modified by Ahn *et al*. (2018) by introducing parallel connections in residual blocks to extract features from multiple levels. Fan *et al*. (2017) proposed a balanced two-stage residual network for image superre solution (BSTRN) which individually retrieved the features of high-resolution and low-resolution images. Yu *et al*. (2018) introduced an SR model which focused on expanding features before applying the activation function Relu, on the residual network. Dargahi *et al*. (2021) proposed a super-resolution method based on a convolutional neural network. In their study, low-resolution images and high-level features are combined in various ways to produce high-resolution images. Proposed an SR model, which reconstructed HR images from noisy images by introducing multiple convolutional blocks for upscaling as well as denoising.

Even though this approach could overcome the vanishing gradient problem through residual blocks, architecture becomes more complex, hence difficult to maintain.

Feedback Network

In the feedback network, a feedback mechanism is introduced to rectify previous stage results and thus generate a more refined output image. Haris *et al*. (2020) introduced the Super-Resolution Feedback propagation Network (SRFBN). In this model, feedback blocks and dense skip connections are introduced. The network's feedback mechanism allows for rectifying faults from higher levels to lower levels and regenerating output images.

Recursive Structures

Recursive networks are non-linear networks in which the same set of weights is applied recursively on structured input to produce structured output. Kim *et al*. (2016b) proposed Deeply Recursive Convolutional Network (DRCN), since it does not converge easily it is modified by Tai *et al*. (2017) by introducing residual blocks along with deep and recursive structures.

Densely Connected Structures

To improve the reconstruction accuracy of the SR model, dense structures are introduced. Tong *et al*. (2017) proposed an SR model using dense skip connections (SRDensenet) which employed dense connections in each layer and across layers so that lower-level features and higher-level features are utilized for image reconstruction. Later this model was modified by Zhang *et al*. (2018) and named as Residual Dense Network (RDN). Even though the model is computationally expensive, it uses dense connections in combination with a residual network to improve reconstruction accuracy by extracting all of the input image's hierarchical features. Kim *et al*. (2021) applied both residual and dense connections in their model to regenerate the high-resolution image. de Leeuw den Bouter *et al*. (2022) applied convolutional network-based SRDenseNet for reconstructing good-quality MRI images. Even though it is a lightweight model, the reconstructed image is upscaled two times only.

Attention Based Networks

Selective attention is included in this SR model to improve the results of chosen regions. A deep convolutional neural network with selection units for super-resolution (SelNet) was introduced by Choi and Kim (2017). The selection unit will determine which feature map from which layer will be passed to the next layer.

Generative Models

Previous models have focused on improving the Peak Signal Noise Ratio (PSNR) between the reconstructed image and the input image. A pleasing appearance of an image will always have good perceptual quality. GAN models focus on minimizing perceptual loss. GAN (Generative adversarial network) has a generator network that learns to generate SR images from lower solution images and a discriminator network that checks the generated image's resemblance to the original image.

A generative adversarial network for Super-Resolution (SRGAN) was described by Ledig *et al*. (2017). In this model, feature extraction is done using a deep residual network with skip connection and a discriminator network with eight convolution layers and dense layers and it focused on three losses: The MSE loss, which measures pixel similarity between low-resolution image and reconstructed image, the perceptual loss, which measures high-level information captured by the generator network and the adversarial loss, which is measured by the discriminator network. Despite the low peak signal-to-noise ratio value, this model achieved better perceptual quality. A major limitation of generative models is the complexity in both generator and discriminator networks Wang *et al*. (2018) proposed.

ESRGAN, which is a modified version of SRGAN. Ahn *et al*. (2022) proposed an SR model with many levels of residual blocks as a generator and applied multiple discriminator blocks for image reconstruction. However, since both the generator and the discriminator must be trained on huge data sets, this model has a high computing need that makes it difficult to maintain in all applications, especially real-time ones.

Auto-Encoder Based SR

Auto-encoders are neural network designs that use encoder networks to learn lower-dimensional feature representations (latent representation) of input images and this latent representation is used by the decoder part to reconstruct an output image of the same dimension. Figure 2 shows the general auto-encoder network diagram. Auto-encoders are suitable for unsupervised learning and are only able to meaningfully compress data similar to the data set on which it is trained. Hence this model is more suitable for all image processing applications to reconstruct a good quality image from its minimal representation.

Pal *et al*. (2018) built an auto-encoder-based SR model for text identification, The encoder in their model consists of two convolution layers with a filter kernel of size 3×3 is used and is then performed up-sampling. The decoder is made up of convolution layers with a filter size of 2×2, with ReLU as the activation function for all except the last layer and sigmoid as the last layer's activation function. But reconstructed image to fade for higher magnification order. Zeng *et al*. (2015) proposed coupled auto-encoders-based SR model. The first auto-encoder acts as encoder part 1 to find an intrinsic representation of the up-scaled version of the LR image and the second

DOI: 10.3844/jcssp.2022.1051.1061

auto-encoder acts as encoder part 2 to find an intrinsic representation of the up-scaled version of the HR image. These intrinsic representations are employed to estimate non-linear mapping from LR to HR and the decoding layer uses mapping as a feedback layer to regenerate the output image. Liu *et al*. (2021) introduced variational auto-encoders for reference-based super resolution. They employed VGG-19 as an encoder for extracting all promising features and the decoder part also contains the VGG network along with bi-linear interpolation for image reconstruction. Andrew *et al*. (2021) employed convolutional auto-encoders with fewer skip connections for image reconstruction on MRI images, model summary shown in Fig. 3 and it is unable to extract all promising features of a low-resolution image. Kebiri *et al*. (2021) employed a convolutional auto-encoder with no skip connections for reconstructing good-quality MRI images.

The review of the SR models stated before makes it evident that there are numerous models for creating high-quality images from low-resolution images. High-quality images in a range of scales must be provided for all image processing applications with the least amount of computing and testing and training complexity possible. Deep learning-based applications' success rates are influenced by the learning approach, the data set's accessibility, and the model's architecture. In these circumstances, auto-encoder-based SR models can reconstruct images with greater PSNR and SSIM values while requiring little extra time for training and testing.

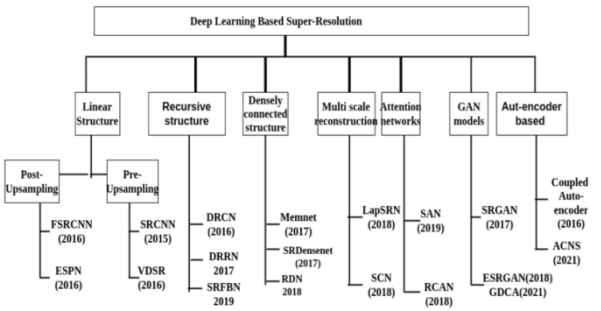

Fig. 1: Major approaches for super-resolution

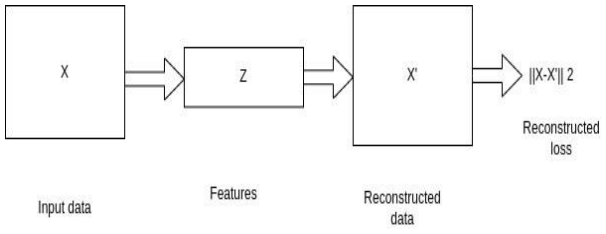

Fig. 2: Auto-encoder network

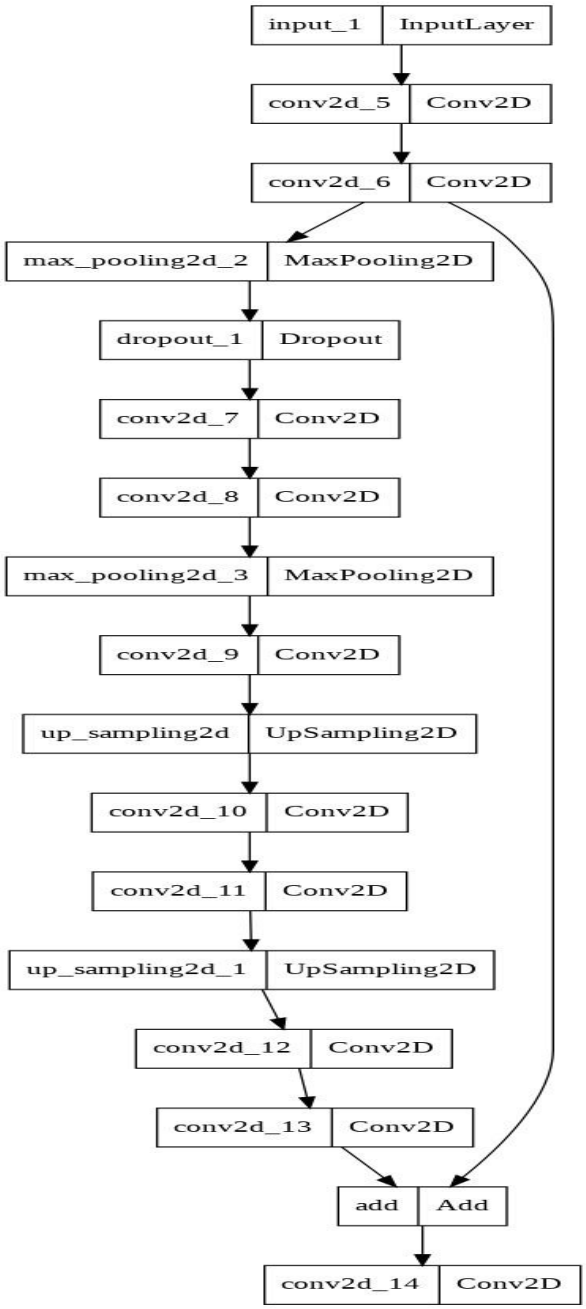

Fig. 3: Previous auto-encoder-based SR model with one skip connection (Andrew *et al*., 2021)

Proposed Auto-Encoder Based SR

Methods

The objective of this study is to generate an HR image from a single LR natural image using neural network architecture. This model applied a convolutional autoencoder with six parallel (local and global) skip connections to extract all prominent features of the input image. We can represent the encoder function as:

$$f1: X_1(\textit{input image}) \rightarrow Z(\textit{Latent representation}) \quad (1)$$

where function *f*1 maps the input image into compressed representation *Z* (*Latent representation*), which represents prominent features used for describing this input image. Encoder network represented as:

$$Z = \sigma(w_1 * X_1 + b) \quad (2)$$

where, * is the convolution operator, αvδ σ is the activation function applied in the network with a bias value '*b*' for filter weight value 'w_1' to train the network. In this model, the relu activation function and the 'adam' optimizing function are applied.

Similarly, the decoder function is given as:

$$f2: Z(\textit{Latent representation}) \rightarrow X_1' \quad (\textit{Regenerated output as similar as input image}) \quad (3)$$

Here decoder network will reconstruct output image X_1', most similar to input image X_1, and the decoder network is represented as:

$$X_1' = \sigma'(w_1' * X_1 + b') \quad (4)$$

and the network is trained to minimize the loss function, which finds the loss of information between the input image and the regenerated output image. It is represented as:

$$L(X_1, X_1') = \|X_1 - X_1'\|^2 \quad (5)$$

The proposed convolutional auto-encoder-based SR model consists of feature extraction, higher-level to lower-level feature mapping, and image reconstruction phases.

Figure 4 shows the proposed model architecture. The natural image applied in this model is of size 256 × 256 and reconstructed image size based on scaling factors. To extract all prominent characteristics from the input image, the encoder uses three of these convolution blocks, followed by a max-pooling process. The decoder consists of convolution blocks and up-sampling operations followed by add operation for combining different layers of high dimensional space and low dimensional space to achieve skip connections and it helps to aggregate both low-level and high-level features extracted by convolution blocks effectively form latent vectors. Convolutional layers with the proper number of filters of size 3 × 3 applied in each layer and "relu" activation functions with "adam" optimizer.

Implementation: Proposed Model Encoder Summary

In this model, the encoder has two CONV2D blocks and operates on an image of 256 × 256 sizes, and applied 64 filters of size 3 × 3. After that max-pooling operation is applied followed by a dropout operation to avoid overfitting. Again, to extract higher-level features, two CONV2D operations are performed with 128 filters and followed by max-pooling and 256 filters applied in the next CONV2D block.

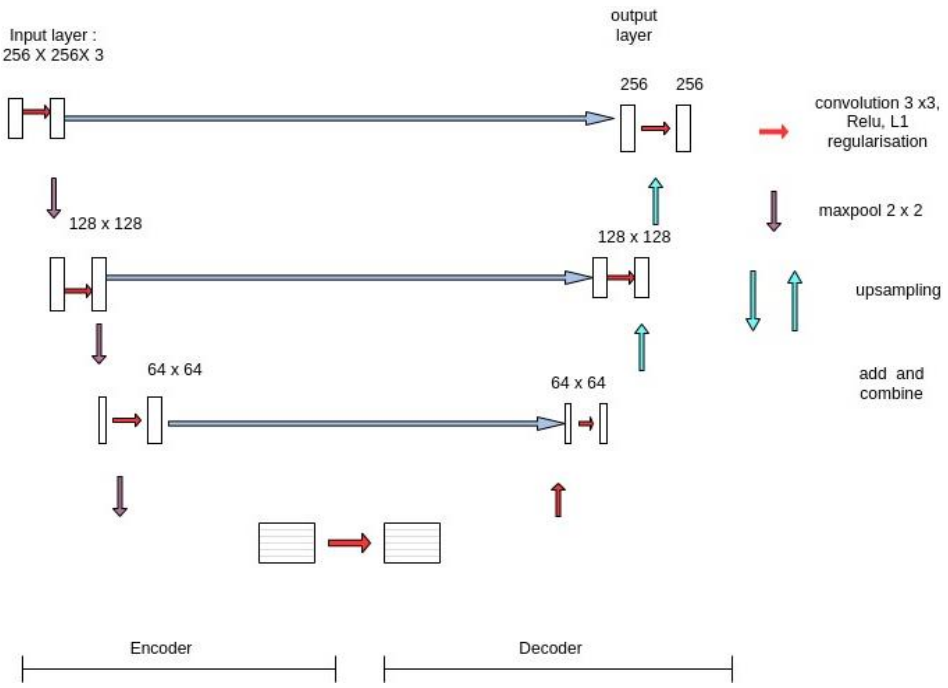

Fig. 4: Architecture of proposed model

DOI: 10.3844/jcssp.2022.1051.1061

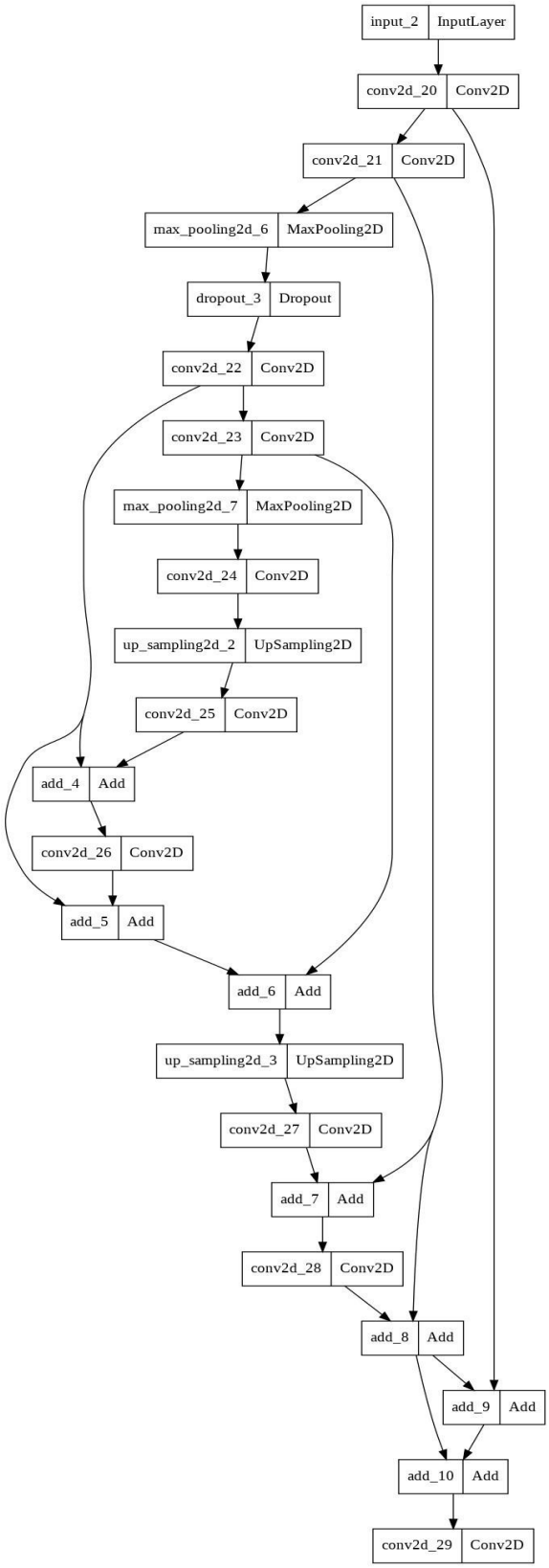

Fig. 5: Proposed model summary with parallel skip connection

Proposed Model Summary-Auto-Encoder with Parallel Skip Connections

In this model, the first convolutional layer extract more generic features like vertical edges, horizontal edges, and diagonal edges, the second layer learns the composition of features learned in the first layer and extracts features like curves and corners and the third layer extracts more domain-specific features. Global and local skip connections are introduced using add and combine operations between initial layers and deeper layers of convolutional blocks to extract all relevant features of the image and convey them to the output layer. The model summary is shown in Fig. 5 here we employed six parallel skip connections for extracting all features effectively and described them as follows. Output from the second convolution layer of the first block in the encoder part is added with a second convolutional layer of the third block in the decoder part for avoiding information loss while reconstruction. The first layer of the third convolution block in the encoder part is added with convolution layers at both the beginning and ending layers of the decoder so that promising features for reconstructing high-resolution images are not lost. End layers of the encoder part are added with the beginning layers of the decoder so that specific features for reconstructing high-resolution images are conveyed to the decoder.

Proposed Model Decoder Summary

The decoder part is composed of three blocks of convolution layers and two up sampling operations in which each convolutional block consists of two convolution operations with a suitable number of filters for adding and combining upper layers with lower layers. Add operation is applied to minimize information loss from the encoder to the decoder in the reconstruction pipeline.

Results and Discussion

Data-set Applied During Test and Training Process

Table 1 shows details of the data set used for testing and training. During testing, sample images from the test data set are employed and evaluated model performance. Fig. 6 shows re-generated image in proposed model and comparison with other methods for test sample images.

Performance Evaluation

Image quality is evaluated objectively using Peak Signal-to-Noise Ratio (PSNR), Mean Square Error (MSE), and Structural Similarity Index (SSIM). This metric is defined as:

$$MSE = \frac{1}{nm}\sum_{1}^{n-1}\sum_{1}^{m-1}\left(x_{j,k} - y_{j,k}\right)^2 \quad (6)$$

DOI: 10.3844/jcssp.2022.1051.1061

where, x and y represent the original image and re-generated image at j and k pixel positions of an $n \times m$ size image.

Peak Signal-to-Noise Ratio (PSNR): Peak Signal-to-Noise Ratio (PSNR): PSNR is measured in dB and has an inverse relationship to mean squared error. It is stated as:

$$PSNR = 20\log_{10} \max f \Big/ \sqrt{mse} \quad (7)$$

where, $max\ f$ denotes the highest frequency value in the input image. The quality of the reconstructed image is indicated by greater PSNR values.

Fig. 6: Test result-comparison with other state of art methods and ground truth image

DOI: 10.3844/jcssp.2022.1051.1061

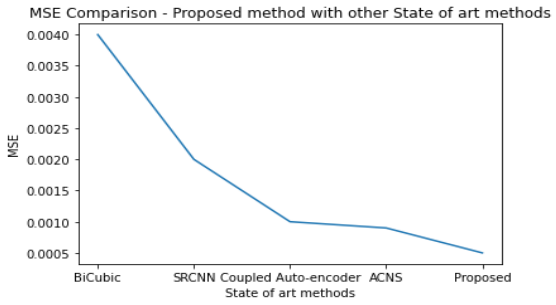

Fig. 7: MSE comparison of the proposed method with state of art methods for test sample images

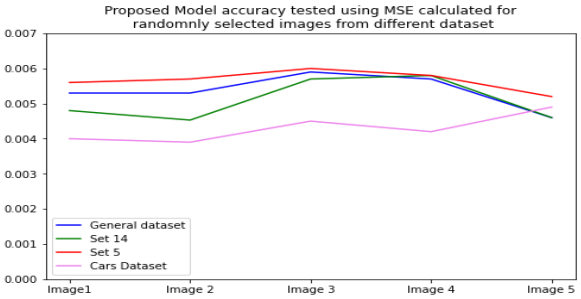

Fig. 8: Mean MSE value comparison of the proposed method with state of art methods on Set5, Set14, Cars, and General data-set

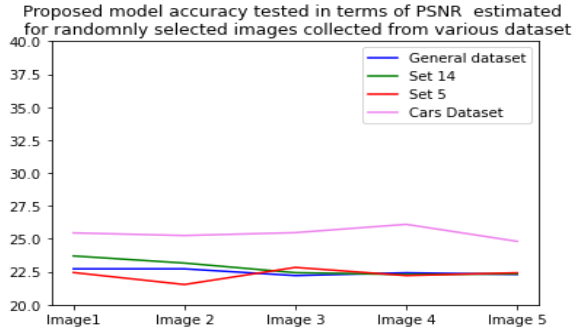

Fig. 9: Mean PSNR comparison of the proposed method and state of art methods using Set5, Set14, Cars, and General data-set

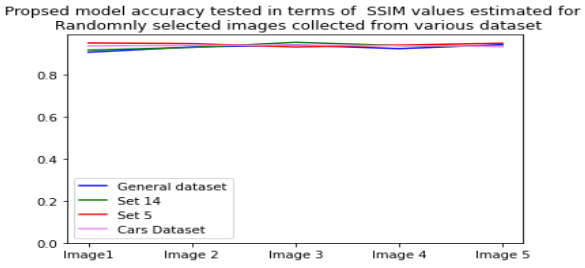

Fig. 10: Mean SSIM comparison of the proposed method and state of art methods using test sample images from Set5, Set14, Cars, and General data-set

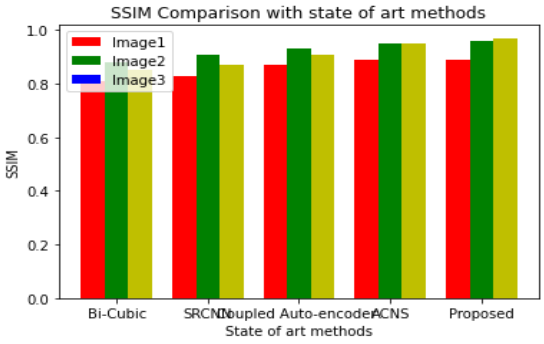

Chart 1: Mean SSM comparison with state of art methods for test sample images from Set5, Set14, Cars Data-set

Structural similarity index: More consistently and accurately than MSE and PSNR, SSIM measures similarity between two images in terms of textures and structure. Given is the distance between two images of size $M \times N$, x, and y. The measure between two images x and y of size $M \times N$ is given as:

$$SSIM(x,y) = \frac{(2 * \bar{x} * \bar{y} + K_1)(2 * \sigma_{xy} + K_2)}{(\sigma_x^2 \sigma_y^2 + K_2) * ((\bar{x})^2 + (\bar{y}))^2 + K_1} \quad (8)$$

where, K_1 and K_2 are constants, $\bar{x}$, $\bar{y}$ represent the average of x and the average of y. σ_x, σ_y represents the standard deviation between ground-truth and regenerated images. SSIM values vary from -1 to 1:

$$\bar{x} = \sum_{1}^{M-1} x_i \quad (9)$$

$$\bar{y} = \sum_{1}^{N-1} y_i \quad (10)$$

$$\sigma_x^2 = 1/M - 1 \sum_{1}^{M-1} x - \bar{x} \quad (11)$$

$$\sigma_{y^2} = 1/N - 1 \sum_{1}^{N-1} y - \bar{y} \quad (12)$$

Performance analysis of the proposed model is done through objective evaluation methods. Here PSNR is compared with the input image and reconstructed image and also PSNR is computed between the low-resolution image and a ground-truth image. Table 2 shows the PSNR value computed for a set of sample low-resolution images concerning the reconstructed image. Similarly, SSIM and MSE values were also computed in this fashion, and seem that these metrics also have a higher value for the reconstructed image than the ground-truth image obtained for SSIM and the least value obtained for MSE. Chart 1 and Fig. 7 show SSIM and MSE results.

Table 1: Data set details

Data set	Number of images and usage	Image size	Source
Div_2 k	634, training	700 × 600	https://data.vision.e e.ethz.ch/cvl/DIV2K/
Cars	8144, training	600 × 400	http://ai.stanford.e du/~jkrause/ cars/car_dataset.h tml
Set5	5, testing	256 × 256	https:/.kaggle.com/datasets/ll01dm/set-514-superresolutiondataset
Set 14	14, testing	256 × 256	https://www.kaggle.com/datasets/ll01dm/set-514-superresolutiondataset
Genera l	100, testing	512 × 512	Kaggle data set

Table 2: Mean PSNR value results on Set5, Set14, general and cars dataset

Data-set	Bi-Cubic	SRCNN	Autoencoder with fewer skip connections	Proposed method
Set5	27.90	28.11	31.21	33.12
Set14	25.87	28.51	30.17	34.87
Cars	25.77	26.33	29.87	33.53
General	24.41	26.87	28.98	32.32
General	24.52	25.35	28.68	33.11

Chart 1 shows SSIM computed using the low-resolution image concerning the reconstructed image and compared results obtained with other state of art methods for a set of sample images. From chart 1, it is clear that for all test cases, the values obtained in the proposed method are higher than in other methods. Figure 7 shows MSE values computed between low-resolution images and reconstructed images for a set of test data-set and comparison with other state of art methods and it is the least value compared to other methods.

Proposed Model Accuracy Analysis

In this study, the consistency of the model was tested using two data sets during training and three data sets during testing time. Figure 8 to 10 shows MSE, PSNR, and SSIM values computed using sample images on this data set. The result shows that the values of evaluation metrics used are not varying much for all sample images. Hence proposed model is better and more consistent than other methods.

Conclusion

The auto-encoder with more parallel skip connections than other existing models employed in this study, which is used to extract all low-level and high-level features from the input image resulting in a single image super-resolution method that is both qualitatively and quantitatively promising. Results compared to other cutting-edge approaches and performance indicators, as well as assessments of our model behavior on various data sets, showed that the suggested strategy outperformed other models.

The average PSNR, SSIM, and MSE values for the proposed technique on several data sets demonstrate the consistency of our suggested model. In all image processing applications especially in real-time applications, miniature scale images are demanded during transmission or storage, but during processing for analysis purposes, it is to be up sampled in different magnification order with no loss of quality is demanded. The performance of this model can be improved by adopting transfer learning concepts in which the encoder trained using a pretrained model is under process and will publish the results soon.

Acknowledgment

Grateful to Dr. Shailesh Sivan (Faculty, CUSAT) for his help and guidance.

Author's Contributions

All authors equally contributed in this study.

Ethics

This study work not published elsewhere and it is our own work.

References

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). http://openaccess.thecvf.com/content_ECCV_2018/html/Namhyuk_Ahn_Fast_Accurate_and_ECCV_2018_paper.html

Ahn, N., Kang, B., & Sohn, K. A. (2022). Efficient deep neural network for photo-realistic image super-resolution. *Pattern Recognition*, *127*, 108649. https://doi.org/10.1016/j.patcog.2022.108649

Andrew, J., Mhatesh, T. S. R., Sebastin, R. D., Sagayam, K. M., Eunice, J., Pomplun, M., & Dang, H. (2021). Super-resolution reconstruction of brain magnetic resonance images via lightweight autoencoder. *Informatics in Medicine Unlocked*, *26*, 100713. https://doi.org/10.1016/j.imu.2021.100713

Choi, J. S., & Kim, M. (2017). A deep convolutional neural network with selection units for super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 154-160).

Dargahi, S., Aghagolzadeh, A., & Ezoji, M. (2021, April). Single Image Super Resolution Using Multi-path Convolutional Neural Network. In *2021 5th International Conference on Pattern Recognition and Image Analysis (IPRIA)* (pp. 1-6). IEEE. https://ieeexplore.ieee.org/abstract/document/9483573

de Leeuw den Bouter, M. L., Ippolito, G., O'Reilly, T. P. A., Remis, R. F., van Gijzen, M. B., & Webb, A. G. (2022). Deep learning-based single image super-resolution for low-field MR brain images. *Scientific Reports*, *12*(1), 1-10. https://www.nature.com/articles/s41598-022-10298-6

Dong, C., Loy, C. C., & Tang, X. (2016, October). Accelerating the super-resolution convolutional neural network. In *European conference on computer vision* (pp. 391-407). Springer, Cham. https://link.springer.com/chapter/10.1007/978-3-319-46475-6_25

Dong, C., Loy, C. C., He, K., & Tang, X. (2015). Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *38*(2), 295-307. https://ieeexplore.ieee.org/abstract/document/7115171

Fan, Y., Shi, H., Yu, J., Liu, D., Han, W., Yu, H., ... & Huang, T. S. (2017). Balanced two-stage residual networks for image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 161-168).

Gilman, A., Bailey, D. G., & Marsland, S. R. (2008, January). Interpolation models for image super-resolution. In *4th IEEE International Symposium on Electronic Design, Test and Applications (delta 2008)* (pp. 55-60). IEEE. https://ieeexplore.ieee.org/abstract/document/4459509

Haris, M., Shakhnarovich, G., & Ukita, N. (2020). Deep back-projectinetworks for single image super-resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *43*(12), 4323-4337. https://doi.org/10.48550/arXiv.1803.02735

Kebiri, H., Canales-Rodriguez, E. J., Lajous, H., De Dumast, P., Girard, G., Aleman-Gomez, Y., ... & Cuadra, M. B. (2021). Through-plane super-resolution with autoencoders in diffusion magnetic resonance imaging of the developing human brain. *bioRxiv*. https://www.biorxiv.org/content/10.1101/2021.12.06.471406v1.abstract

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

Kim, J., Lee, J. K., & Lee, K. M. (2016b). Deeply-recursive convolutional network for image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1637-1645).

Kim, S., Jun, D., Kim, B. G., Lee, H., & Rhee, E. (2021). Single image super-resolution method using CNN-based lightweight neural networks. *Applied Sciences*, *11*(3), 1092. https://doi.org/10.3390/app11031092

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

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

Lim, B., Son, S., Kim, H., Nah, S., & Mu Lee, K. (2017). Enhanced deep residual networks for single image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 136-144). https://doi.org/10.48550/arXiv.1707.02921.

Liu, Z. S., Siu, W. C., & Wang, L. W. (2021). Variational autoencoder for reference based image super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 516-525).

Nasrollahi, K., & Moeslund, T. B. (2014). Super-resolution: A comprehensive survey. *Machine Vision and Applications*, *25*(6), 1423-1468. https://link.springer.com/article/10.1007/s00138-014-0623-4

Pal, S., Jana, S., & Parekh, R. (2018, December). Super-resolution of textual images using autoencoders for text identification. In *2018 IEEE Applied Signal Processing Conference (ASPCON)* (pp. 153-157). IEEE. https://ieeexplore.ieee.org/abstract/document/8748679

Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A. P., Bishop, R., ... & Wang, Z. (2016). 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* (pp. 1874-1883). https://doi.org/10.48550/arXiv.1609.05158

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).

Tong, T., Li, G., Liu, X., & Gao, Q. (2017). Image super-resolution using dense skip connections. In *Proceedings of the IEEE international conference on computer vision* (pp. 4799-4807). https://ieeexplore.ieee.org/abstract/document/9289414

Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., ... & Change Loy, C. (2018). Esrgan: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European conference on computer vision (ECCV) workshops* (pp. 0-0). http://arxiv.org/abs/1809.00219

Yu, J., Fan, Y., Yang, J., Xu, N., Wang, Z., Wang, X., & Huang, T. (2018). Wide activation for efficient and accurate image super-resolution. *arXiv preprint arXiv:1808.08718*.https://arxiv.org/abs/1808.08718

Zeng, K., Yu, J., Wang, R., Li, C., & Tao, D. (2015). Coupled deep autoencoder for single image super-resolution. *IEEE Transactions on Cybernetics*, *47*(1), 27-37. https://ieeexplore.ieee.org/abstract/document/7339460

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).