A Low Complexity Patch Based Approach to Image Super-Resolution for Higher Scale Factors

Joel Alex John *Rutwik Kumar M* *E. Jayanth* *Noel Tony* *Ravindra Kumar Singh*

joelalexjohn9@gmail.com rutwikala@gmail.com ejaynt434@gmail.com noelnoel255@gmail.com singh.ravindrakumar@gmail.com

Department of Electronics and Communication Engineering

Christ University Bangalore, India

Abstract: Single Image Super Resolution (SISR) is the process of mapping a low-resolution image to a high-resolution image. This has proven to be useful in a wide range of applications including enhancement of image quality of medical images, CCTV footage, or satellite images. Recent developments in deep learning have shown significant improvements in SISR. While these methods are effective, they come with the cost of higher complexity. Therefore, in our paper we have proposed a low complexity approach to increase the quality of an image. A key contribution of our work is addressing a research gap of not considering higher scale factors as prior works concentrate more on x2 and x4 scale factors. We have used a patch-based approach to preprocess images as it is easier to recover high frequency components of the low-resolution image by following this approach. Rather than using an architecture with increased depth which can increase the complexity of the model, we have used Residual Blocks which can also help in preserving important features. Furthermore, we have incorporated a Pixel Attention mechanism to find out the pixels with higher importance. Experimental results show that our model can perform with higher efficiency and lesser complexity on images with a higher level of degradation.

1. Introduction

In the present day, media files are often compressed by applications to reduce their size for easier sharing, but this efficiency comes at the cost of diminished quality, resulting in a reduced quality experience for the end user. The single-image super-resolution method involves training models to upscale and improve the detail of a single low-resolution (LR) image to a high-resolution (HR) image. The process also holds a lot of real-life applications such as Medical Diagnosis [17], remote sensing [18] and intelligent surveillance [19]. The image can be degraded due to various factors such as blur, decimation, or noise, or it can simply be due to the traditional bicubic interpolation method. [5].

The field has seen considerable progress recently, largely due to the evolution of Convolutional Neural Network (CNN). A lot of different architectures have been used to serve the purpose. Chao Dong et al (2015) have developed a Super Resolution model using a Convolutional Neural Network (SRCNN), which is one of the earliest attempts at resolving the problem. It has just three convolutional layers that map between low- resolution and high-resolution images [21].

The Enhanced Deep Super-Resolution network (EDSR), is a streamlined version of the traditional ResNet model, designed by removing redundant layers to make it more compact. Despite its simplified structure, EDSR still achieves great performance with state-of-the-art results.[5]. On the other hand, one can always take a different approach, particularly a generative approach to resolve this issue, Baozhong Liu et al (2021). have come up with a Super Resolution model with a Generative Adversarial Network (GAN). The model uses a generator network that produces pseudo-high-resolution images and a discriminator network that evaluates the difference between generated and real high-resolution images. Although the Peak signal-to-noise ratio (PSNR) value of generative models aren't appealing, the visual results of the model are satisfactory [22].

While these models excel in specific aspects, they come with limitations. Many super-res olution techniques, including EDSR and SRCNN models struggle while attempting larger scaling factors [5][21], like x6, where maintaining finer details become increasingly difficult, often leading to visual artifacts, resulting in blurred textures and distorted structures that degrade the overall image quality and models like GAN are computationally demanding, requiring significant processing power and memory. This makes them challenging to train efficiently and impractical for deployment in low-resource or real-time applications.

This paper discusses about a patch-based method using the Efficient Sub-Pixel Convolution Network (ESPCN) model. The model was designed by Shi et al (2016). for image super-resolution with a focus on efficiency and performance [16]. It is an updated version of the traditional CNN model with an added sub-pixel Convolutional layer to the network [5]. The proposed ESPCN model is relatively faster than other models. However, the original ESPCN model also is known to struggle with upscaling the images at a higher scaling factor [16]. The base model is capable of delivering better results when the patch-based method is used. In the context of image processing, Patching refers to a technique where the image is broken down into smaller sections or "patches" where it has quite a few advantages as it allows the models to understand the finer details better. The patching has been done so that the stride size is half the size of each patch, which leads to overlapping patches, resulting in most parts of the image being processed twice, which leads to improved contextual understanding and introduces redundancy [11]. We have also incorporated an additional feature known as the Pixel Attention Network. The attention mechanism evaluates the important pixels and removes the pooling operations that aggregate information. This allows the model to keep only the necessary details and reduces the number of parameters [6].

2. Related Works

There have been many works done prior to this in the field of image super-resolution. Various techniques, from traditional methods to architectures using deep learning, have been explored for this.

Traditional methods for image super-resolution used interpolation techniques such as bicubic and bilinear interpolation [23], which work by estimating the pixel values between existing pixels. However, since it assumes that pixel values change smoothly, it cannot extract high-frequency details like textures or sharp edges. Later on, more advanced methods, such as sparse coding [24] and dictionary learning [25], were introduced.

The advent of deep learning [26] significantly advanced the field of super-resolution. SRCNN [27] model pioneered convolutional neural networks (CNNs) [28] for super-resolution by using deep networks to produce high-resolution images from low-resolution images. Since then, numerous architectures have been proposed to enhance further performance, including deeper networks like VDSR [9], EDSR [14], and SRGAN [29]. Kim, J et al. (2016) developed a series of deep neural networks that learned to map low-resolution features to high-resolution features through its layers. Lim, W et al. (2017) developed another architecture that has improved upon the VDSR [9] architecture. EDSR [14] has residual blocks, which help the model learn more complex details. Finally, SRGAN [29] is another advancement in image super-resolution with the help of deep learning. SRGAN, developed by Ledig, C et al. (2017), is a type of GAN [30] in which the generator tries to generate a high-resolution output of a low-resolution input image. However, all these models lacked in one way or another. Most deep learning architectures were complex and computationally expensive. Some models could even distort the input image, which made things worse.

Lightweight architectures have been explored to improve computational efficiency and real-time applicability. The Efficient Sub-Pixel Convolutional Neural Network (ESPCN) [16] introduced a sub-pixel convolution layer to upscale images, reducing the computational complexity [16] and still having satisfactory outputs. Due to its reduced complexity, it focuses on learning efficient upscaling filters and has become a baseline for real-time SR applications.

Patch-based super-resolution [21] handles low-resolution images more effectively in computer vision applications. Patch Extraction focuses on extracting smaller square patches of size (PxP) from the image, processing them independently and then concatenating to form the final output.

Pixel Attention [6] is another mechanism used for image super-resolution. X. Wang et al. (2020) proposed this method which reconstructs a high-resolution image using weighted pixels. Each pixel is assigned a weight based on its relevance, and how apposite they are while considering key details like texture and edges.

An acute way to improve the performance of a deep learning model is using skip connections [31]. Proposed by Srivastava, R et al. (2015), this is a technique that connects some layers to their previous layers, thereby combining features from both levels. This preserves the information from earlier stages, hence extracting more features. A type of skip connection that is functional for image super-resolution is residual blocks [32]. He, K et al. (2016) proposed this technique with a series of convolutional layers followed by a skip connection.

Our approach builds upon the ESPCN model [16] by incorporating a patch-based preprocessing method. Instead of reshaping the entire image to fit the model input, we first create uniform square patches of the image. The network then processes each patch independently, allowing the model to handle arbitrary image shapes without resizing or distorting the input. The processed patches are then reassembled to form the final high-resolution output. This method retains the advantages of the sub-pixel convolution introduced in ESPCN while improving performance on non-square images and reducing memory overhead for high-resolution inputs.

3. Methodology

3.1 Data Preprocessing

Images in the dataset has been normalized to be between 0 and 1. It has also been observed that the model has better performance when the training images are in the YUV format. This is because the YUV image format allows the model to focus more on the luminance of the image from which its quality can be determined [3]. The low-quality images were created by downscaling and upscaling the images by a factor of 6 using bilinear interpolation.

3.2 Patch Creation

Due to the similarity between adjacent pixels in a patch [2] lost information can be recovered easily. For a smaller image, the model only has to recover less amount of high frequency components compared to the entire image. It can also be observed that the use of typical interpolation techniques to up sample images assumes uniformity in the image and smooth transitions and estimates new pixel values by averaging the nearby values [4]. This can lead to the creation of images with high artifacts and could not capture the high-frequency details in the image. However, a patch-based approach can capture the underlying meaning of patterns in an image [4]. This was also done to create overlapping regions within the patches to extract more features [3].

A set of patches were extracted from each of the images from the training dataset. The patch size is kept as 10 pixels wide while the stride has a value of 5. Hence an input image I of shape H x W x C is turned into P x P x C.

3.3 Model Architecture

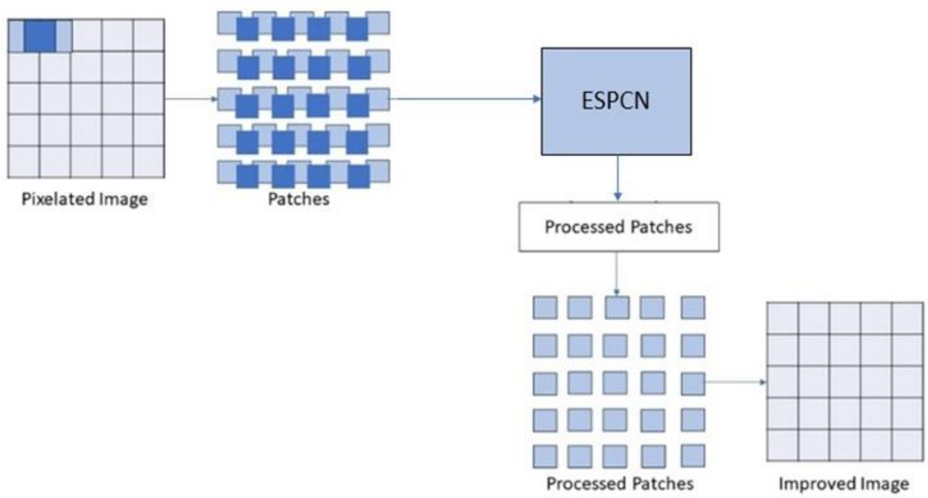


Fig. 1

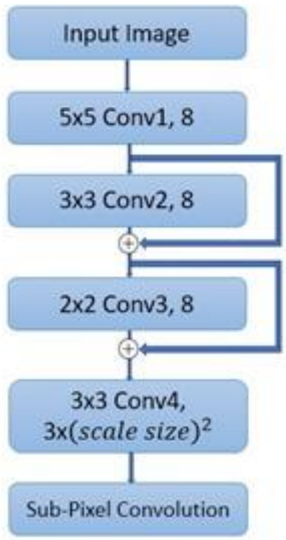


Fig. 2

3.3.1 Base Model

An Efficient Sub-Pixel Convolutional Network (ESPCN) has been used here and has been tuned for a patch-based super resolution approach. The model has four main convolutional blocks with the first three convolutional layers having 16 filters each.

To effectively capture a broader range of features, a 5x5 kernel was used at the first layer. After this the feature map generated is then passed into a 3x3 convolutional layer and then to a 2x2 convolutional layer to extract finer details from the image.

All the layers used a custom activation function which was a combination of Leaky ReLU and linear function to help in retaining identity mapping.

Y = LeakyReLU(x) + x

where x is the input to the node.

3.3.2 Residual Blocks

Residual blocks have been employed with promising results for various computer vision applications. This is due to their ability to preserve identity information throughout the super resolution process [5]. These skip connections can help information from earlier layers to be sent to deeper layers without much loss.

Given an input patch P_i with shape P x P x C:

First Convolutional Layer:

$$C_1 = LeakyRelu(W_1 \cdot P_i + b_1)$$

Where W_1 and b_1 are the weights and biases of the first convolutional layer and $*$ denotes convolution.

Second Convolutional Layer with Skip Connection:

$$C_2 = LeakyReLU(W_2 \cdot C_1 + b_2)$$

$$S_1 = C_1 + C_2$$

Third Convolutional Layer with Custom Activation and Pixel Attention:

$$C_3 = X^{attn}\left(LeakyReLU(W3 \cdot S1 + b3)\right)$$

$$S_2 = S_1 + C_3$$

Final Convolutional Layer:

$$C_4 = LeakyReLU(W_4 \cdot S_2 + b_4)$$

$$O = DepthToSpace(C_4, s)$$

Where s is the scaling factor.

3.3.3 Single Pixel Attention

We have introduced the Pixel Attention mechanism proposed by Zhao H. et al (2020) into our model. Attention mechanisms are useful to identify important regions of an image [12], when combined with convolutional layers this can help extract features from key regions of an image. This was done by removing the pooling operation in Channel Attention [7] and Spatial Attention [8] which has shown to significantly improve the performance of the model [6]. The weights given for the pixels help in giving context about the region as a whole [13]. This has been implemented since not all pixels in an image contribute equal amount of semantic information to an image [13].

In the pixel attention block for our model, the feature map from previous layers is passed in as a 4D tensor of the form (B, P, P, F).

$$X \in R^{B \cdot P \cdot P \cdot F}$$

Where B is the batch size, P is the height and width of an image patch and F is the feature map. The average of this is taken in order to reduce its dimension. By averaging, we can calculate the overall importance of pixels giving a 3D tensor (B, P, P). It is important to note that the average has been calculated over the channel or feature axis, which helps in concentrating on preserving the spatial information.

$M_{i,j} = 1/F \; \Sigma X i, j, k$

where $X_{i,j,k}$ is the max activation value

$M_{i,j}$ is the mean value of pixels

The next step is to find the pixel with the highest value which can be used to find the significance in the information provided by each pixel. By using this selective attention mechanism, the model can focus on the regions of the image which are more relevant.

$argmax(i,j)\; M_{i,j}$

After this, we have created an attention mask for that pixel, ignoring the rest of the pixels in that feature map. This helps in removing any pixel which does not contribute much to the model. This was done by encoding the indices of the pixels that contribute more information. The result of this process is a tensor with the pixel that should receive attention while the rest are masked out.

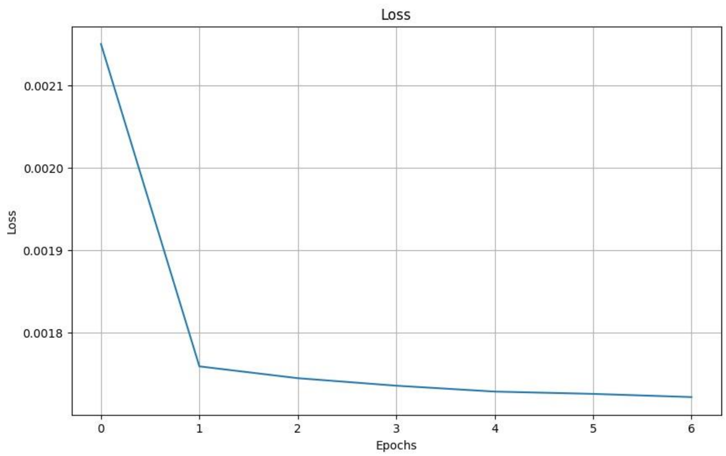
4. Results

Dataset	Scale	VDSR PSNR/SSIM [9]	SRCNN PSNR/SSIM [10]	Bicubic PSNR/SSIM [10]	MDSR PSNR/SSIM [10]	EDSR PSNR/SSIM [10]	ESPCN With patches (Ours) PSNR/SSIM
Urban 100	×2	30.76/0.9140	29.50/0.8946	26.68/0.8403	32.84/0.9347	32.93/0.9351	34.7331/0.9028
	×4	25.18/0.7524	24.52/0.7221	23.14/0.6577	26.67/0.8041	26.64/0.8033	32.6002/0.8116
	×6	-/-	-/-	-/-	-/-	-/-	31.8930/0.7846
B100	×2	31.90/0.8960	31.36/0.8879	29.56/0.8431	32.29/0.9007	32.32/0.9013	39.748/0.9431
	×4	27.29/0.7251	26.90/0.7101	25.96/0.6675	27.72/0.7418	27.71/0.7420	38.6734/0.9100
	×6	-/-	-/-	-/-	-/-	-/-	37.5822/0.8891
SET 5	×2	37.53/0.9587	36.66/0.9542	33.66/0.9299	38.11/0.9602	38.11/0.9601	32.0859/0.7974
	×4	31.35/0.8838	30.48/0.8628	28.42/0.8104	32.50/0.8973	32.46/0.8968	30.2986/0.7088
	×6	-/-	-/-	-/-	-/-	-/-	31.7536/0.7520
SET 14	×2	33.03/0.9124	32.42/0.9063	30.24/0.8688	33.85/0.9198	32.92/0.9195	36.4188/0.9424
	×4	28.01/0.7674	27.49/0.7503	26.00/0.7027	28.72/0.7857	28.80/0.7876	35.3347/0.8990
	×6	-/-	-/-	-/-	-/-	-/-	35.2389/0.8905
DIV 2K Validation	×2	33.66/0.9625	33.05/0.9581	31.01/0.9393	34.96/0.9692	35.03/0.9695	37.0468/0.9313
	×4	28.17/0.8841	27.78/0.8753	26.66/0.8521	29.26/0.9016	29.25/0.9017	34.7141/0.8533
	×6	-/-	-/-	-/-	-/-	-/-	34.7511/0.8327

Table 1: It presents the results of our model ESPCN++, which was assessed using many publicly benchmarked datasets. The performance metrics for VDSR, SRCNN, Bicubic, MDSR and EDSR were obtained for their base models across the datasets Urban 100, B100, Set 5, Set 14 and DIV2K Validation set.

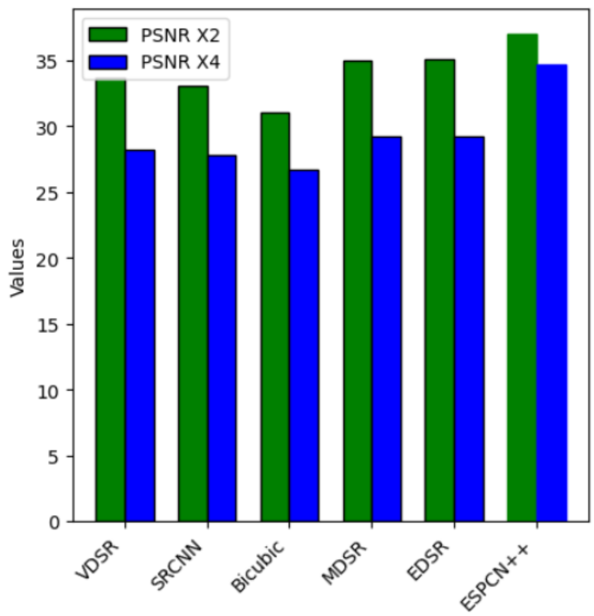
We provide the quantitative evaluation of results of Our final Model (ESPCN++) on the public benchmark datasets in the Table 1. We trained our Model individually on the images and other settings are same as the baseline Models. We compare our Model with state-of-the-art models like VDSR [Original Reference], SRCNN [Original Reference], MDSR [Original Reference]. For comparison we measure PSNR and SSIM on the final RGB output image produced by the Model. We used python function for evaluation. Our Model exhibits significant improvements as compared to other models. The proposed models restore the LR image back to HR version successfully along with the detailed textures and edges in the HR images and produce better-looking SR outputs compared with the previous works.

Through experimentation we found out that dividing the image into patches has led to better performance compared to baseline models. This improvement is because creating patches helps to extract more features than processing a full-sized image. These low-level features, which contain minute details, contribute to a more accurate reconstruction of the high-resolution image. The extraction of these low-level features enables the model to effectively capture finer details that are often lost with full-sized images. Additionally, processing patches is more efficient, reducing training time and computational cost.

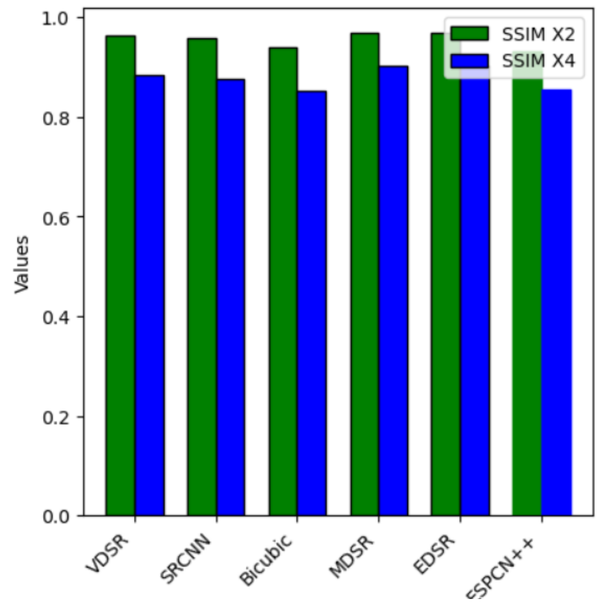


Graph 3: A plot of Mean Square Error (MSE) loss function of the model's performance

The MSE (Mean Squared Error) drops significantly with increase in Epochs, suggesting high conversance of the Algorithm. The limited number of parameters for average of X6 image super resolution the loss is minimized to as low as 0.0004 in just 6 epochs.



Graph 3 : Comparison of PSNR Values across techniques For scaling Factor of x2 and x4



Graph 4: Comparison of Structural Similarity Index (SSIM) values across various techniques for scaling factors of 2x and 4x on the DIV2K validation set.

On comparison with other models our model has produced superior results on DIV2K validation set on the PSNR and SSIM metrics. Our Model easily averages 36 PSNR on 2X and around 34 on 4X scaling factor.

When applied for Super-resolution of images scaled by a factor of 6, the total number of trainable parameters in the model are 20,236. In contrast, the parameters in EDSR are 43.6 million (for X2) [14], for VDSR parameters vary around 66500[15] this can change with different configurations. MDSR has a total of nearly 1.3 million parameters. Our model outperforms these models by a large margin in terms of efficiency and computational cost, making it particularly suitable for deployment in systems requiring fast results.

Models	Number of Parameters
EDSR [14]	43.6 million (for X2)
VDSR [9]	66500
MDSR [15]	1.3 million
Our Model	20236 (for X6)

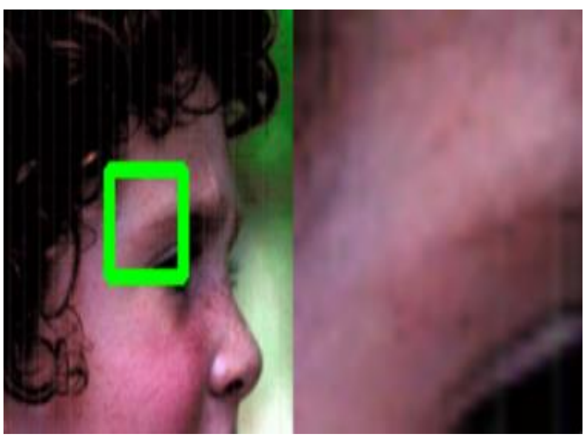
Table 2 :

6X

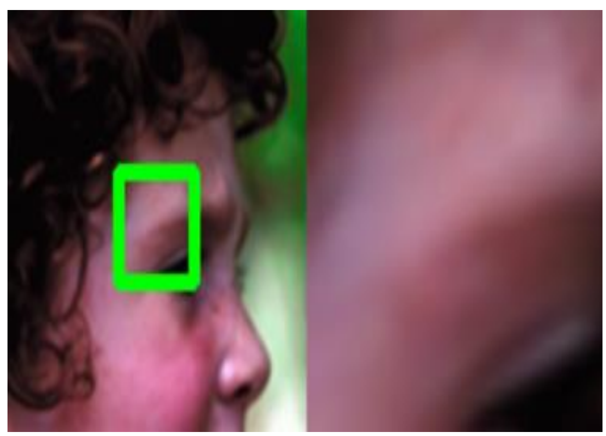
Fig 2. Original Image from set 5 Dataset

Fig 3. Downscaled Image by a factor of 6

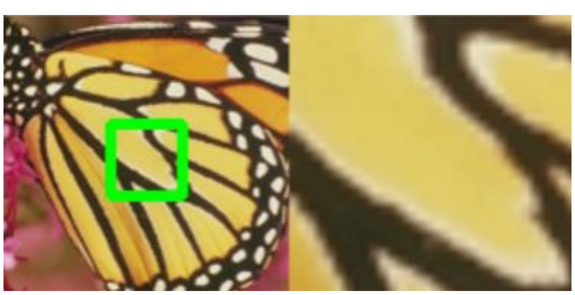
Fig 4. Image after Super Resolution by our Model



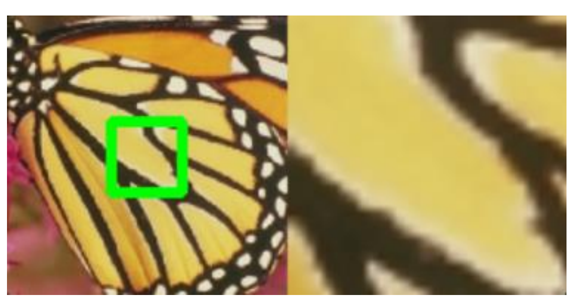
Our Model



EDSR



(a)



(b)

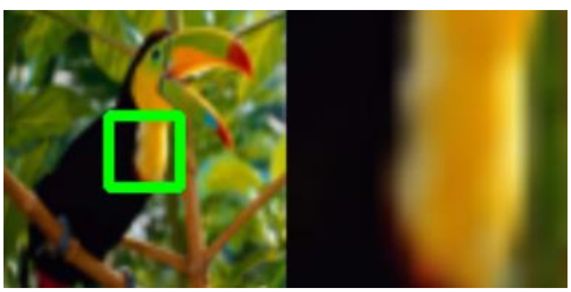
(c)

(d)

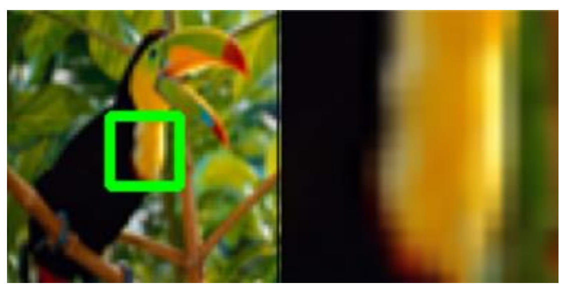
(e)

(f)

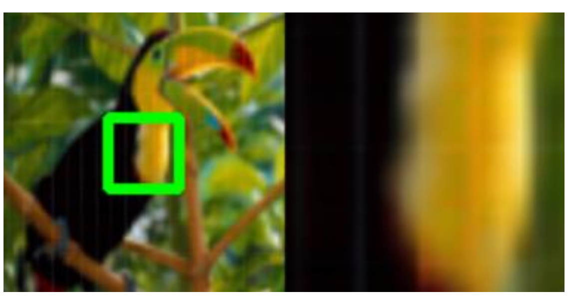
Fig 5: Single image super-resolution results of “butterfly” from Set5 dataset with upscaling factor 2. (a) Ground-truth. (b) SRCNN. (c) Bicubic (d) Our's ESCPN++. (e) EDSR. (f) LapSRN



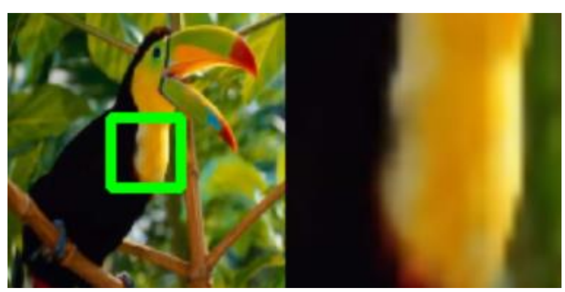
(a)



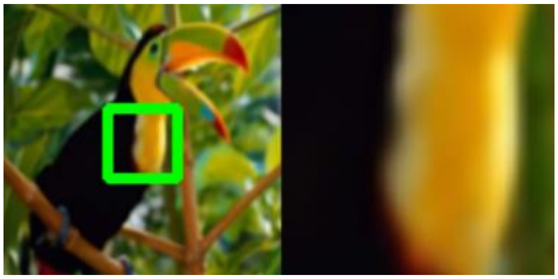
(b)



(c)



(d)



(e)

(f)

Fig 7: Single image super-resolution results of “Parrot” from Set5 dataset with upscaling factor 2. (a) Ground-truth downscaled by a Scaling Factor of 4. (b) Bicubic. (c) Our Model (ESPCN++) (d) SRCNN.
(e) EDSR. (f) LapSRN

In all the Test Images from Fig 5, Fig. 6, Fig. 7 Our Model has performed qualitatively better than the compared Model

5. Conclusion

In this paper we propose an efficient super resolution technique. The main idea behind our architecture is to divide images into patches and feed these individual patches to the base ESPCN Model. This approach enables the model to extract more features from the patches as compared to feeding the whole image at once. Our experiments have exhibited that on employing this procedure, the model outperforms the most of the SISR (Single Image Super-Resolution) models. Our Model accomplishes top ranking results for x2, x4, x6 scaling factor. The super resolution of images with a scaling factor of x6 is the novel findings we have concluded in our paper. Our proposed model surpasses all the current models in super resolution of images with X6 scaling factor. Furthermore, our model is lightweight, with a maximum size of less than 80KB, enhancing its potability across different computing environments. Building on this, future researchers can explore methods for the model to be trained efficiently on higher scale factors such as x8 and x10.

References:

[1] C. Dong, C. C. Loy, K. He, and X. Tang, “Learning a deep convolutional network for image super-resolution,” in ECCV, 2014.

[2] Mikaeli, E., Aghagolzadeh, A. & Azghani, M. Single-image super-resolution via patch-based and group-based local smoothness modeling. *Vis Comput* **36**, 1573–1589 (2020). https://doi.org/10.1007/s00371-019-01756-w

[3] Jiang, Junjun & Hu, Ruimin & Han, Zhen & Lu, Tao. (2014). Efficient single image super-resolution via graph-constrained least squares regression. Multimedia Tools and Applications. 72. 10.1007/s11042-013-1567-9.

[4] Coupé P, Manjón JV, Chamberland M, Descoteaux M, Hiba B. Collaborative patch-based super-resolution for diffusion-weighted images. Neuroimage. 2013 Dec;83:245-61. doi: 10.1016/j.neuroimage.2013.06.030. Epub 2013 Jun 19. PMID: 23791914.

[5] B. Lim, S. Son, H. Kim, S. Nah and K. M. Lee, "Enhanced Deep Residual Networks for Single Image Super-Resolution," 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 2017, pp. 1132-1140, doi: 10.1109/CVPRW.2017.151. keywords: {Training;Image resolution;Computer architecture;Computational modeling;Signal resolution;Image reconstruction;Convolution},

[6] Zhao, H., Kong, X., He, J., Qiao, Y., Dong, C. (2020). Efficient Image Super-Resolution Using Pixel Attention. In: Bartoli, A., Fusiello, A. (eds) Computer Vision – ECCV 2020 Workshops. ECCV 2020. Lecture Notes in Computer Science(), vol 12537. Springer, Cham. https://doi.org/10.1007/978-3-030-67070-2_3

[7] Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the

IEEE conference on computer vision and pattern recognition. pp. 7132{7141 (2018)

[8] Woo, S., Park, J., Lee, J.Y., So Kweon, I.: Cbam: Convolutional block attention

module. In: Proceedings of the European conference on computer vision (ECCV).

pp. 3{19 (2018)

[9] Kim, J., Lee, J. K., & Lee, K. M. (2016). Accurate image super-resolution using very deep convolutional networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(10), 2009-2020

[10] Lim, B., Son, S., Kim, H., Nah, S., & Lee, K. M. (2017). *Enhanced Deep Residual Networks for Single Image Super-Resolution*. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops* (pp. 1132-1140). Seoul National University, Seoul, Korea.

[11] Coupé P, Manjón JV, Chamberland M, Descoteaux M, Hiba B. Collaborative patch-based super-resolution for diffusion-weighted images. Neuroimage. 2013 Dec;83:245-61. doi: 10.1016/j.neuroimage.2013.06.030. Epub 2013 Jun 19. PMID: 23791914.

[12] W. Sun and B. -D. Liu, "ESinGAN: Enhanced Single-Image GAN Using Pixel Attention Mechanism for Image Super-Resolution," 2020 15th IEEE International Conference on Signal Processing (ICSP), Beijing, China, 2020, pp. 181-186, doi: 10.1109/ICSP48669.2020.9320934. keywords: {Training;Superresolution;Image resolution;Generators;Generative adversarial networks;PSNR;Task analysis;SinGAN;Image Super-resolution;Pixel Attention},

[13] N. Liu, J. Han and M. -H. Yang, "PiCANet: Pixel-Wise Contextual Attention Learning for Accurate Saliency Detection," in IEEE Transactions on Image Processing, vol. 29, pp. 6438-6451, 2020, doi: 10.1109/TIP.2020.2988568.

[14] Lim, B., Nah, S., & Lee, K. M. (2017). *Enhanced Deep Residual Networks for Single Image Super-Resolution*. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Link

[15] 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 (CVPR). Link

[16] W. Shi et al., "Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 1874-1883, doi: 10.1109/CVPR.2016.207. keywords: {Image resolution;Convolution;Neural networks;Image reconstruction;Interpolation;Feature extraction;Convolutional codes},

[17] Y. Li, B. Sixou, F. Peyrin, A Review of the Deep Learning Methods for Medical Images Super Resolution Problems, IRBM, Volume 42, Issue 2, 2021, Pages 120-133, ISSN 1959-0318, https://doi.org/10.1016/j.irbm.2020.08.004

[18] D. Yang, Z. Li, Y. Xia and Z. Chen, "Remote sensing image super-resolution: Challenges and approaches," 2015 IEEE International Conference on Digital Signal Processing (DSP), Singapore, 2015, pp. 196-200, doi: 10.1109/ICDSP.2015.7251858. keywords: {Image reconstruction;Remote sensing;Interpolation;Spatial resolution;Signal resolution;Wavelet transforms;Super resolution;remote sensing;observation model},

keywords: {Feature extraction;Saliency detection;Object detection;Context modeling;Convolution;Semantics;Computational modeling;saliency detection;attention network;global context;local context;semantic segmentation;object detection},

[19] Hijji, M.; Khan, A.; Alwakeel, M.M.; Harrabi, R.; Aradah, F.; Cheikh, F.A.; Sajjad, M.; Muhammad, K. Intelligent Image Super-Resolution for Vehicle License Plate in Surveillance Applications. *Mathematics* **2023**, *11*, 892. https://doi.org/10.3390/math11040892

[20] D. Capel and A. Zisserman, "Computer vision applied to super resolution," in *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 75-86, May 2003, doi: 10.1109/MSP.2003.1203211

keywords: {Computer vision;Strontium;Layout;Cameras;Signal resolution;Image resolution;Image restoration;Robustness;Optical imaging;Signal restoration},

[21] C. Dong, C. C. Loy, K. He and X. Tang, "Image Super-Resolution Using Deep Convolutional Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 2, pp. 295-307, 1 Feb. 2016, doi: 10.1109/TPAMI.2015.2439281.

keywords: {Image resolution;Neural networks;Image reconstruction;Convolutional codes;Feature extraction;Training;Super-resolution;deep convolutional neural networks;sparse coding;Super-resolution;deep convolutional neural networks;sparse coding},

[22] B. Liu and J. Chen, "A Super Resolution Algorithm Based on Attention Mechanism and SRGAN Network," in IEEE Access, vol. 9, pp. 139138-139145, 2021, doi: 10.1109/ACCESS.2021.3100069.

keywords: {Superresolution;Feature extraction;Convolution;Image reconstruction;Training;Image coding;Signal resolution;Super resolution;GAN network;attention mechanism},

[23] Wang, Y., Zhan, D. C., & Zhang, Y. (2020). "Towards enhancing the robustness of neural networks by principled networks." arXiv. https://arxiv.org/abs/2009.09622

[24] Olshausen, B. A., Field, D. J. (1997). Sparse coding with an overcomplete basis set: A framework for understanding visual cortex. Vision Research, 37, 3311-3334.

[25] Aharon, M., Elad, M., Bruckstein, A. (2006). K-SVD: "An algorithm for designing an overcomplete dictionary for sparse representation." IEEE Transactions on Signal Processing, 54, 4311-4322.

[26] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). "Learning representations by back-propagating errors." Nature, 323(6088), 533-536. https://doi.org/10.1038/323533a0

[27] Dong, C., Change, C. C., Yang, Y., Li, X. (2014). "Learning a deep convolutional network for image super-resolution." IEEE Transactions on Pattern Analysis and Machine Intelligence, 38, 1005-1017.

[28] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). “Gradient-based learning applied to document recognition.” Proceedings of the IEEE, 86(11), 2278-2324. https://doi.org/10.1109/5.726791

[29] Ledig, C., Theis, L., Huszár, F., Caballero, J., Aitken, A., Tejani, A., Totty, J., Bishop, N., Shi, G. (2017). A learned representation for image super-resolution. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41, 367-380.

[30] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial networks.

arXiv preprint arXiv:1406.2661.

[31] Srivastava, R. K., Greff, K., & Schmidhuber, J. (2015). Highway networks. arXiv preprint arXiv:1505.00387.

[32] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[33]