

Xception

by Wazir Muhammad

Submission date: 05-Oct-2021 11:20PM (UTC+0500)

Submission ID: 1666134750

File name: 4Version_Electronics.docx (1.8M)

Word count: 10068

Character count: 54905

3

1 Article

2 **Multi-scale Xception Based Depthwise separable
3 convolution for Single Image Super-Resolution**4 Wazir Muhammad¹ and Supavadee Aramvith *5 ¹ wazir.laghari@gmail.com6 * Correspondence: Supavadee.a@chula.ac.th; Tel.: +66-2-218-6911.

7 Received: date; Accepted: date; Published: date

8 **Abstract:** Recently, research on deep convolutional neural networks (CNNs) based single image
9 super-resolution (SISR) have been achieved remarkable performance in the area of image, video and
10 computer vision tasks. In spite of their great performance, the network size in terms of depth and
11 width of SISR increases every day, which demands the larger computational resources, more
12 memory consumption and not suitable for real-world applications like desktops, laptops, mobiles,
13 smartphone's, tablets and Internet of Things (IoT) devices. In this paper, we present a depthwise
14 separable convolution using Xception module to learn multi-scale information and design a very
15 deep network for SISR. In addition, the skip connection enables to create short paths and to reuse
16 the feature maps to alleviate the vanishing-gradient problem of the very deeper network
17 architecture. The transposed convolution layer is an integrated before the Xception block to increase
18 the computational efficiency and speed up the training process. Moreover, at the later end, we used
19 Xception block to recover the high quality and high-resolution (HR) image. Our proposed approach
20 does not used any interpolation technique as the preprocessing step and thus dramatically decrease
21 the size of the model in terms of number of parameters, while increasing the visual representation
22 efficiency of the model. Furthermore, most deep learning based SISR approaches used Rectified
23 Linear Unit (ReLU) as a non-linear activation function, but problem with ReLU have a gradient
24 constant in most cases, so to solve this issue we interchange the ReLU function with Parametric
25 Rectified Linear Unit (PReLU) activation function. PReLU is not only above advantage, but also can
26 make faster convergence speed and to avoid the dead features. Additionally, we adapt the Adam
27 optimizer which can increase the training process and work well on LR image. Compared with
28 existing state-of-the-art algorithms, our proposed algorithm has demonstrated a great ability to
29 construct HR images with fine, rich and sharp texture details as well as edges. Additionally, we
30 assess our model using images from five public benchmark datasets and compare with recent deep
31 CNN based methods. The experimental results validate that the our proposed approach has robust
32 performance compared to other popular existing approaches in terms of accuracy, speed, and visual
33 quality.

34

35 **Keywords:** image super-resolution; convolutional neural network; depthwise separable
36 convolution; xception module.

37

19 1. Introduction

38 Single image super-resolution (SISR) aims to obtain the visually attractive high-resolution (HR)
39 output image from a single input low-resolution (LR) image generating by low cost imaging
40 framework within the limited environmental conditions. Recently, SISR is a very interesting research
41 space in the area of image and computer vision tasks, which is extensively applied in various
42 applications, such as an object detection [1, 2], image segmentation [3, 4], and image classification [5,

45

43 ¹ purposes. The better performance and higher accuracy of SISR have been encouraged in the field
44 ¹ of medical imaging [7-9], face detection and recognition [10, 11], high-definition television (HDTV)
45 [12], video surveillance [13], satellite imaging [14] and autonomous driving technology [15, 16],
46 where rich details information are greatly desired.

47 Recently, many SISR based algorithms tends to develop for better performance and to
48 reconstruct the high quality visual pleasing results. Generally, image SR approaches can be
49 categorized in to main three classes [17]: Interpolation-based methods [18-20], reconstruction-based
50 methods [21] and learning-based methods [22-32].

51 Interpolation-based image SR methods, such as bilinear interpolation, nearest neighbor
52 interpolation, bicubic interpolation and Lanczos interpolation [18, 33, 34]. These approaches are
53 simple and implementation is very easy, but still, they suffer from accuracy shortcomings and
54 generating the jagged ringing artifacts, as well as produce the blurry results in the smooth region
55 areas.

56 Reconstruction based image super-resolution methods [21, 25, 31, 35-37] mostly adopted
57 previous information to restrict the feasible solution with an advantage of reconstructing the sharp
58 ¹ details and suppressing the statistical noise effects [38]. However, these methods are mostly time
59 consuming and image reconstruction performance degrades rapidly on 4x or 8x scale factor
60 enlargements.

61 Learning-based image SR methods are brought into focus by researcher due to outstanding
62 ²⁹ performance and fast computation. These methods usually apply the machine learning approaches
63 to evaluate the relationships between the LR input image and its corresponding counterpart HR
64 output image from the training samples. Chang et al. proposed the idea of neighbor embedding [29]
65 to take the benefit of similar patches generated by locally to reconstruct the HR image from LR input
66 image patches. Researchers used the idea of sparse signal recovery theory [39], and introduces the
67 concept of sparse coding methods [31, 32, 40-43] to solve the SISR problem. Meanwhile,
68 reconstruction-based methods combine with learning-based methods to further reduce the jagged
69 ringing artifacts and improve the blurry results [44-47].

70 Currently, deep neural networks [48-50] provide significantly improved performance and
71 led to dramatic changing in SISR. Furthermore, deep neural networks approaches are very fast and
72 accurate, but still there is some limitations. However, existing deep convolutional neural network
73 model stacked the convolution layer side by side to create the deeper network architecture, due to
74 this increase the computational cost and introduce the vanishing gradient problem during the
75 training. In addition with, existing deep convolutional neural network approaches used bicubic
76 interpolation technique as a pre-processing step to upscale the LR image and to introduce the new
77 noises in the model, because interpolation technique does not designed for this purpose. In order to
78 solve such issues and further enhance the image quality of existing methods, we proposed the novel
79 multi-scale Xception based depthwise separable convolution for single image super-resolution
80 (MXDSIR) to reconstruct the visually pleasing high quality HR image from low quality LR image.

81 In summary, our main contributions through this paper are three folds:

- 82 • We proposed the residual convolution block and replace the standard convolution with
83 depthwise separable convolution to achieve faster convergence in the training and to avoid the
84 vanishing gradient problem as well as ease the training complexity.

- 85 • We replace the Rectified Linear Unit (ReLU) with Parametric Rectified Linear Unit (PReLU) to
86 activate the dead features that is caused by zero gradients in the ReLU activation function.
87 • We introduced the new Xception block, which can detect the different image features
88 information to reconstruct the high quality image.

89 The remainder of this paper is organized as follows. In the Section 2 we review the literature
90 work of image super-resolution methods. Section 3 describes the proposed network architecture of
91 single image super-resolution in detail. Our proposed network architecture and its experimental
92 evaluations with other existing state-of-the-art approaches are explained in the Section 4, and Section 5
93 finally conclude the work with opinions and discussions.

94 2. Related Works

95 The target of SISR is to construct the visually pleasing, high quality and high resolution output
96 image that has more details information from the original input LR image. However, SR is a
97 challenging ill posed problem. There are three techniques can be used to solve the image SR problem.
98 The first technique is the interpolation based technique, which is easy in implementation and very
99 simple, but cannot obtain the high frequency details. Reconstruction based is the second technique,
100 that is applied for smoothing and downsampling to recover the high resolution image. For rare
101 situations, the reconstructed images have a lot of blurring and artifact noise. The third and most
102 successful technique is the example based approaches that uses machine learning based techniques
103 to construct the HR images. As this technique is growing rapidly and many other methods have been
104 developed, such as the sparsity-based method [31], using sparse coding to learn dictionaries of low
105 and high resolution images. For regression-based techniques [40] that uses the regression type
106 approaches for faster and better image reconstruction. After the popularity of deep CNN learning
107 based architecture, many researcher starts to implement deep learning based convolutional neural
108 network techniques to solve SISR problem. Various approaches have been suggested to solve SISR
109 problem, but here we discuss only on recent deep convolutional neural network based approaches.

110 The first concrete deep learning based approach for SISR problem proposed by Dong et al. [51],
111 known as Super-Resolution Convolutional Neural Network (SRCNN) and reported significant
112 improvements over all previous SR methods. SRCNN model used three convolution layers to predict
113 the output mapping function from upscaled bicubic version of the image to reconstruct the HR image.
114 Although, there are some weakness of this model. First, the proposed approach used the bicubic
115 interpolation as pre-processing step to upscale the original LR image, because bicubic interpolation
116 is introduced the blurry results and not designed for this purpose. Second, reconstruction of low,
117 middle and high level feature information is still unsatisfactory. Third convergence of training is
118 very slow, and take more time.

119 Dong et al. [52] further improved SRCNN [51] method and proposed the faster version of an
120 hourglass-shaped architecture to accelerate super-resolution image reconstruction, named as Fast
121 Super-Resolution Convolutional Neural Network (FSRCNN) [52]. FSRCNN has a modest network
122 architecture, that depends on four CNN and one deconvolution layer for upsampling purpose and
123 used the original input LR images without interpolation techniques. FSRCNN have lower
124 computational complexity and better performance as compared to SRCNN, but has a limited network
125 capacity.

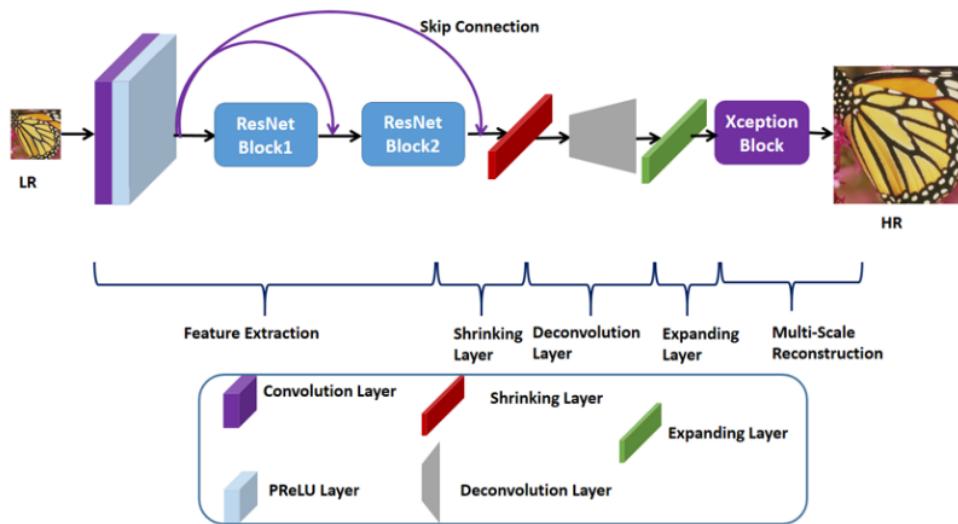
126 ¹⁴ Efficient sub-pixel convolutional neural network (ESPCN) [53] is a simple, efficient and fast
127 image SR method, that can apply on real time image and video applications. ESPCN reduce the
128 computational complexity to replace the bicubic up-sampling operation with sub-pixel convolutional
129 layer to extracts the rich features information from original low resolution image. Kim et al. [48]
130 proposed an accurate image SR method by using a very deep convolutional network (VDSR) inspired
131 by Visual Geometry Group network (VGG-net) used for ImageNet classification purpose [5]. VDSR
132 network reported the significant performance and improvement over the SRCNN network by
133 increasing the depth of the network up to 20 CNN trainable layers. To ease the training complexity
134 of a deeper model uses the global residual learning with a fast convergence speed, to calculate the
135 difference between original ground truth image and bicubic interpolation upsampled version of LR
136 image. However, VDSR network architecture does not used the actual pixel values, but used the
137 bicubic interpolated based image as an input, that leads to heavier computation cost and more
138 memory consumption. Wang et al. [54] introduce the sparse prior deep convolutional neural
139 networks for image super resolution based approach, named as Sparse Coding Network (SCN). This
140 approach is easy, simple and achieve remarkable performance over SRCNN. Wang et al. further
141 modified the model and replace the non-linear mapping by a set of sparse coding sub-networks [55].
142 The major drawback of SCN is the high computational complexity and also hinders its applications
143 in real-time processing scenarios. Kim et al. proposed Deeply-Recursive Convolutional Network for
144 image super-resolution (DRCN) [49] and uses the convolution layers multiple times. The key
145 advantage of DRCN is to fixed the number of training parameters, although there are more number
146 of recursions, but main deficiency is to slow the training process. Authors also used recursive type
147 supervision and skip connections to optimize the performance of the model. Very deep Residual
148 Encoder-Decoder Networks (RED) [56] used symmetric convolution and deconvolution layers with
149 residual learning. RED model is trained on 30 layers and achieves the best performance. Therefore,
150 these studies reflect the concept of “the Deeper the Better”.

151 Contrarily, a shallow and fast deep learning based approach proposed by Yaniv et al. named as
152 Rapid and Accurate Image Super Resolution (RAISR) [57]. In this approach author classifies the input
153 image patches with respect to angle of patches, coherence and strength to learn the maps from LR
154 image to HR. Lei et al. [58] suggested a deep Laplacian Pyramid Super-Resolution (LapSRN) to
155 reconstruct the HR image. LapSRN architecture depends on the multiple pyramid levels and each
156 level is caused by a deconvolution layer as an upsample, but problem is that it uses a fixed integer
157 scaling factor, that is limited the flexibility of the network. Zhang et al. [59] proposed the feed forward
158 denoising convolutional neural networks (DnCNNs) to accelerate the improvement of a very deep
159 convolutional neural network architecture to enhance perceptual quality of the image. DnCNN
160 follow the same architecture as SRCNN and stacked the convolution layer followed by batch
161 normalization with ReLU activation function. Although model provide the favorable results but they
162 are computationally expensive due to the use of batch normalization layer. Zhao t al. [60] propose a
163 more flexible scaling factors to super-resolve the input LR image named as gradual upsampling
164 network (GUN). GUN uses the bicubic interpolation technique to upsample the features and the
165 gradients in the forward and the backward computation during the training. Tai et al. [61] introduced
166 the idea of deep recursive residual network (DRRN) to identify the number of model parameters,
167 which recursively learns the residual unit in a multi-path model. Ledig et al. [50] employ a deep

168 residual network connections with 16 blocks by skip-connection to recover the LR image with an 1
 169 upscaling factor of 4x. Lim et al. [62] proposed an enhanced deep super-resolution network to 1
 170 increase the training efficiency of the model by removing the batch normalization layers and their 1
 171 method win the NTIRE2017 SR challenge [63]. Meanwhile, Tai et al. [64] suggested the deepest 1
 172 persistent memory network (MemNet) for image restoration, in which a multiple memory blocks are 1
 173 stacked to obtain the persistent memory. Yamanaka et al. [65] proposed a combined architecture of 1
 174 skip connection layers and parallelized CNN layers to develop a deep CNN architecture for image 2
 175 super-resolution. This model is shallower than VDSR. Deep CNN with Skip Connection and Network 1
 176 in Network (DCSCN) [65] model uses the different networks, such as extraction of different levels of 1
 177 features network and SR image reconstruction type network. W.M et al. proposed multi-scale 1
 178 inception based super-resolution using deep learning approach (MSISR) [66] for image 1
 179 reconstruction. In this approach authors uses the concept of asymmetric convolution operation to 1
 180 reduce the complexity of the model and finally used the inception block to reconstruct the multiscale 1
 181 feature information for image SR. Li et al. used an adaptive feature detection process to obtained the 2
 182 features fusion at different scales, named as multi-scale residual network (MSRN) [67]. This approach 1
 183 used the complete hierarchical type of feature information to reconstruct the accurate image super- 1
 184 resolution.

185 3. Proposed Methodology

186 In this section, we discuss our proposed network architecture for SISR based on ResNet blocks 1
 187 and Xception based block. As similar to the existing SISR methods, our proposed method can be sub- 1
 188 divided into five stages: feature extraction, shrinking, up-sampling, expanding and multi-scale 2
 189 reconstruction stages. Figure 1, presents the overall network architecture of our proposed model. 1
 190 The main target of SISR is to reconstruct an image Y^* , which is the upscaled form of LR image X , 1
 191 where we want Y^* to be the same as the ground-truth image Y .



192

193

1 Figure 1. Proposed Network Architecture of Xception Based Single Image Super-resolution reconstruction.

194

3.1. Feature Extraction

This part is similar to the previous methods, but different on the input image. However, most previous deep convolutional neural network types SISR approaches are used to extract the features from a bicubic interpolated version of HR image. Actually, bicubic interpolation is to damage important information of LR image and introduce new noises in the model, which may play key role in recovering the HR image [68]. Furthermore, extracting the feature information of up-sampled version of pixels are redundant, especially, in the case of enlargement factor of more than 3x [65]. Therefore, our approach used an alternate technique, to extract the features information directly from LR image without interpolation technique and network can hold the feature information efficiently. Our initial feature extraction part/stage consists of one convolution layer and two ResNet Blocks with skip connection to extract the low, middle and high level features information simultaneously. Inspired by VDSR, we used the one convolution layer of filter size 3×3 with 64 number of filters followed by Parametric Rectified Linear Unit (ReLU). Mathematically convolution layer can be explained as:

208

$$F_l = \max(0, W_l * F_{l-1} + b_l) \quad (1)$$

, where W_l is the weights of the filter and b_l is the biases of the l^{th} layers, respectively; the output of the feature map denoted by F_l and "*" represents the convolution operation. The W_l support $n_l \times f_l \times f_l$ number of parameters, f_l indicates the filter size, n_l represent the number of filters. The convolution layer and ResNet blocks have the same sizes of $3 \times 3 \times c$, kernels to generate the c features map, where c represents the numbers of image channels or 64 channels.

214

3.1.1 PReLU

Earlier approaches uses the convolution layers or blocks followed by the Rectified Linear Unit (ReLU) [69], like SRCNN and VDSR. These types of model have fair response, but still results are not satisfactory, because in most cases ReLU [69] gradient is a constant and to avoid the vanishing gradient problem. In the proposed model we used Parametric Rectified Linear Unit (PReLU) [70], which is not only solve the said problem, but also has faster speed of convergence during the training. Mathematically, PReLU [70] activation function can be explained as:

221

$$PReLU(x_i) = \max(x_i, 0) + a_i \min(0, x_i), \quad (2)$$

222

, where x_i denoted activation function of the input signal on the i^{th} layer, and the PReLU coefficient of the negative part is denoted by a_i , such a_i parameter is used to set zero value for ReLU [69] activation function, but it is for PReLU [70] for learnable parameter purpose. The main purpose of PReLU [70] is used to avoid the "dead features", which is caused by zero gradient in the ReLU [69] activation function. Finally, the output feature maps with PReLU activation function can be written as:

229

$$F_l(Y) = PReLU(W_l * F_{l-1}(Y) + B_l) \quad (3)$$

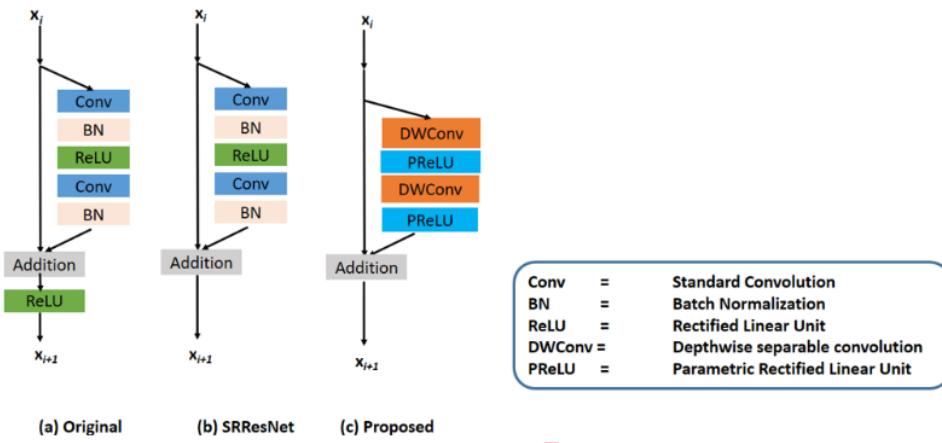
230

, where F_l denoted the resultant output features map, B_l denoted the biases of l^{th} layer.

232

3.1.2 Feature Extraction Blocks

233 The layer stacked side by side to enhance the network depth, in such a way deeper network
 234 architecture has weaker information received at the final layers [71], which introduce the vanishing
 235 gradient problem and increases the computational cost of the model [6]. He et al. [72] proposed the
 236 ResNet blocks to resolve the above said problems and extensively used in the deep learning type SISR
 237 image to reconstruct the HR image. The ResNet blocks [48, 50, 59, 72] have developed as a family of
 238 deeper network type architectures, which shows the high accuracy and faster convergence behaviors
 239 in the area of machine learning, computer vision and speech recognition. Furthermore, the deeper
 240 ResNet architecture has been effectively used in the field of SISR and demonstrated superior
 241 performance [50, 62]. In our proposed method uses a different residual skip connections which makes
 242 the faster training and reduces complexity of the model.



243 (a) Original (b) SRResNet (c) Proposed

244 Figure 2. The comparisons diagrams of different residual blocks with proposed ResNet block.
 245 (a) Original ResNet block. (b) SRResNet without final ReLU activation function. (c) Our Proposed
 246 ResNet block, remove the BN and replaces the standard convolution and ReLU activation function
 247 with depthwise separable convolution followed by PReLU activation function.

248 In Figure 2, we shows the comparison diagrams of original residual skip [72] connection,
 249 SRResNet [50] and our proposed ResNet block. The architecture of the ResNet block as shown in
 250 Figure 2(a) uses a direct path with skip connection for transmitting the features information through
 251 residual skip connection, and summed up the resultant information followed by ReLU activation
 252 function. SRResNet [50] block as shown in the Figure 2(b), uses the alternative strategey to remove
 253 the ReLU activation function and provide the simple and clean path from one block to another block.
 254 Figure 2(c), shows our proposed ResNet block, which eliminates the Batch Normalization (BN) [73]
 255 layers to reduce the usage of Graphics Processing Unit (GPU) memory and enhance the
 256 computational efficiency of the model. Furthermore, we replace the operation of standard
 257 convolution with depthwise separable convolution followed by point wise convolution to reduce the
 258 number of parameters and ReLU activation function replaced with PReLU. The PReLU is used to
 259 avoid the vanishing gradient problem and to reduce the training complexity as well as increases the
 260 efficiency of the model. For middle and high level feature extraction, we applied 2 ResNet blocks,
 261

262 each block consists of two 3×3 depthwise separable convolution kernel with 64 filters followed by
263 PReLU non linearity.

264 3.2 Shrinking Layer

265 If more number of features are directly apply on the transpose convolution layer, which will
266 create high computational cost and increase the size of the model. Thus, it is very important to reduce
267 the input feature maps to keep the model compactness as well as to reduce computational cost of the
268 model. It has been proved in earlier study [74] that a 1×1 convolution layer used as a shrinking or
269 bottleneck layer to decrease the size of the input feature maps. To take the benefit of 1×1 convolution
270 layer, we employ a 1×1 convolution layer as a shrinking layer before the deconvolution layer. This
271 type of arrangement is also observed in latest deep convolutional neural network types models for
272 computer vision tasks. Authors in [66, 68, 75] used the shrinking layer to increase the computational
273 efficiency of the model.

274 3.3 Deconvolution Layer

275 In [55, 76, 77] researchers suggested that increase the LR image resolution before the first layer
276 is to increase the computational cost as well as damage very important information, because
277 processing speed is directly depends on the resolution of the image. Furthermore, interpolation
278 technique used before the first layer, do not provide additional information to solve the image
279 reconstruction ill-posed problem. To handle these types of problems, we proposed to reconstruct the
280 HR image directly from original LR feature domain. For this purpose we apply the deconvolution
281 layer as an upscaling operation before the Xception block, with small 3×3 filters of stride is equal to
282 enlargement scale factors to transforms the feature mapping information from LR to HR space.

283 3.4 Expanding Layer

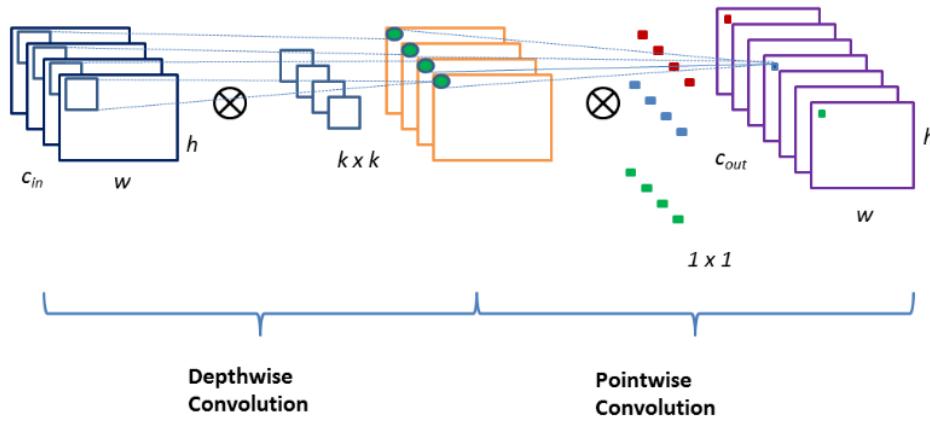
284 The expanding layer performs the inverse operation of shrinking layer. The main operation of
285 shrinking layer is to reduce the dimensions of LR features for the purpose of computational efficiency.
286 However, if reconstruct the directly HR image from LR features, the final reconstruction quality of
287 the image will be poor. Therefore, we apply the inverse shrinking layer or expanding layer after the
288 deconvolution to recover the original features information back.

289 3.5 Multi-Scale Reconstruction

290 At the final stage, we used a multi-scale Xception block stands for multi-scale Extreme version of
291 Inception block, which is adopted from GoogLeNet [78] with a modified depthwise separable
292 convolution better than Inception v-3 [74], which is also proposed by GoogLeNet for image
293 classification tasks. Multi-scale Xception block is used to select the appropriate kernel size, because
294 kernel size plays a very important role in the model design, training procedure as well as multi-scale
295 reconstruction of SR image. Generally, smaller size is better for capturing the information locally, and
296 the larger kernel size is more preferable for information distributed globally. The Xception network
297 uses this idea and include many depthwise separable convolution with a different kernel size like
298 3×3 , 5×5 and 7×7 .

299 3.5.1 Depthwise Separable Convolution

300 Originally depthwise separable convolution was proposed by Laurent Sifre [79] and was applied
 301 for image classification purpose. The depthwise separable convolution is a type of factorize
 302 convolution, which performs the factorization operation of a standard convolution, which convert
 303 into a depthwise convolution followed by pointwise convolution operation. In the Figure
 304 3, the depthwise separable convolution operation performs a single filter per input channel, and to
 305 combine the linearly input channels. The convolution operation replaces with a two-layer factorized
 306 convolution layer, one is used for space filter and other is used for combining purpose. Hence
 307 depthwise separable convolution can significantly reduce the number of parameters as well as size
 308 of the model.



310 Figure 3. Basic diagram of depthwise separable convolution followed by a pointwise convolution
 311 operation.

312 The standard type of convolution kernel takes three parameters, such as height (h), width (w)
 313 and input channel (c_{in}) of an input feature map (I), the resultant convolution layer is $(h \times w \times c_{in})$ is
 314 applies as $K \times K \times c_{in} \times c_{out}$ convolutional kernel J to produce an output feature map of O as $h \times w \times$
 315 c_{out} , where c_{out} is the number of output channels as shown in the Figure 3. The depthwise separable
 316 convolution depends on two convolution operations: depthwise separable convolution operation and
 317 pointwise convolution operation. We uses depthwise separable convolution and apply a separate
 318 one filter for each input channel of feature maps. Mathematically, the depthwise separable
 319 convolution operation [80] can be written as:

$$320 \quad G(y, x, j) = \sum_{u=1}^k \sum_{v=1}^k K(u, v, j) \times I(y + u - 1, x + v, j) \quad (4)$$

321 ,where K represents the kernels of depthwise separable convolution operation of size $k \times k \times c_{in}$. The
 322 n^{th} filter in the kernel K is applied on the p_{th} number of channels in the input feature map of I to
 323 reconstruct the G output feature map of n^{th} number of channels. To reconstruct the new features we
 324 apply the pointwise convolution. Mathematically, pointwise convolution can be written as:

$$325 \quad O(y, x, l) = \sum_{j=1}^{c_{in}} G(x, y, j) \times P(j, l) \quad (5)$$

326 ,where the size of kernel of pointwise convolution operation is $1 \times 1 \times c_{in} \times c_{out}$.

3.5.2 Xception Blcok

At the final stage, we used a Xception block [78], which is adopted from GoogLeNet [74] to select the appropriate size of the kernel. Kernel size performs an important job in the network, as well as to stabilize the training procedure of the model, because size of kernel is very close connection to extract a suitable features information. Smaller kernel size is the best choice for locally captured the information, and globally distributed information is captured by the larger kernel size. This concept is used by the Xception network and involves several highly separable convolutions with a different kernel sizes. Furthermore, the inception architecture of second and third version uses the same idea of standard convolution operation, which has high computational cost as compared depthwise separable convolution.

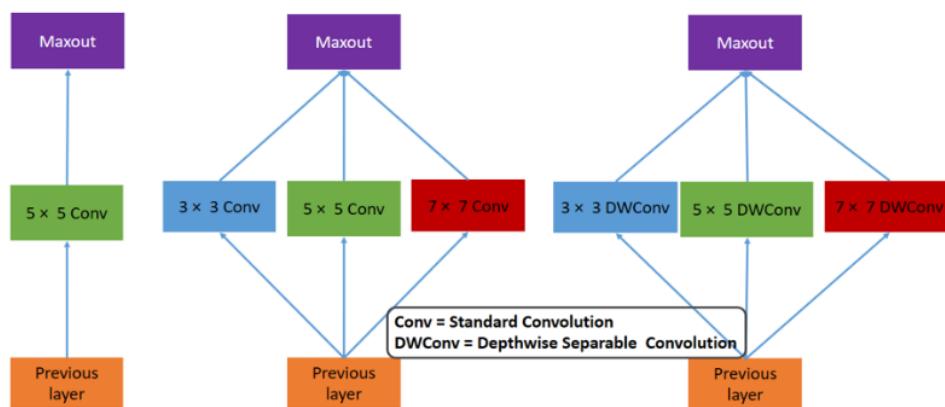


Figure 4. Comparison of single and multi-scale standard and depthwise separable convolution blocks [53]. (a) Single Scale Convolution (b), Multi-Scale Standard Convolution, and (c) Multi-Scale Depthwise Separable Convolution (Our proposed).

In Figure 4 (a), shows a single scale plain convolution type architecture, in which several convolution layers are stacked in a single straight line path, such type of architecture is implemented by well known image super-resolution methods, like SRCNN and FSRCNN architectures. These types of architectures are simple, but a deeper network architecture increases the depth of the model and more memory consumption. In Figure 4 (b), uses the standard convolution type inception block to extract the multi-scale feature information efficiently. Though, such type of blocks has more number of parameters and create the overfitting problem during the training. In Figure 4(c), shows our proposed block of multi-scale depthwise separable convolution, to solve the over and under fitting training problems as well as reduce the computational cost of the model. Proposed Xception block consists of different depthwise separable convolution kernel size, like 3×3 , 5×5 and 7×7 , followed by point wise convolution with PReLU activation function, to reconstruct the SR image.

353 *Experimental Results*

In the experimental section, first of all we explain the selection procedure of training datasets with different model hyper-parameters. Afterward, we evaluate the quantitative as well qualitative performance in terms of PSNR/SSIM and perceptual vision quality on publicly five benchmark test

357 datasets. Finally, we compares the computational cost, processing speed of our proposed model in
 358 terms of PSNR versus number of K parameters and running time.

359 3.2. Training Datasets

360 There have been various sizes of image datasets for training purpose to train the model for single
 361 image-super resolution, but Yang et al. [31] and the Berkeley Segmentation Dataset (BSDS100) [81]
 362 are most commonly used image datasets, because these datasets are used by well-known SR methods.
 363 like VDSR, DRCN and LapSRN for training purpose. To evaluate the proposed method, we take 91
 364 images from Yang et al. [31] and another 200 images from BSDS300 datasets [81]. As followed-by [82],
 365 and to take the advantage of training dataset , stabilizing the training process, to avoid the problem
 366 of over-fitting and enhancing the capacity of the model, we used the data augmentation technique in
 367 terms of rotation and flipping.

368 All experimental evaluations were done on the original ground truth image. For Data
 369 manipulation purpose, we used a programming language developed by MathWorks (MatLab 2018a),
 370 and deep learning Keras 2.2.1 library supported back-end as TensorFlow. Various types of loss
 371 functions were available to validate the performance of our proposed model. Mostly deep learning
 372 type convolutional neural networks based SR architecture have used the loss function as a mean
 373 squared error (MSE), so we also used the similar type of loss with our proposed method.
 374 Mathematically, loss function can be calculated as:

$$375 \quad L(\theta) = \frac{1}{m} \sum_{i=1}^m \|F(Y_i; \theta) - X_i\|^2 \quad (5)$$

376 ,where $F(Y, \Theta)$ is the final recovered image, X_i is the high-quality HR images, Y_i is the
 377 corresponding low quality or LR image, and m is the samples numbers of small sizes batches during
 378 the training.

379 In the training phase we used adaptive momentum estimation optimizer (Adam) [83] having
 380 0.0003 initial learning rate with mini-batch size is 32. The training process takes 200 epochs to
 381 converge model properly. All experimental results were conducted on an Ubuntu 18.04 operating
 382 system with NVIDIA Titan Xp GPU, having 3.5 GHz Intel i7-5960x CPU with 64 GB RAM. We trained
 383 our model on luminous channel to speed up the training procedure.

384 3.3. Testing Datasets

385 We evaluate the performance of our proposed network architecture on five benchmark test
 386 datasets. The Set5 [84] dataset consists of five images having different sizes like 228 x 228 and 512 x
 387 512 pixels. The Set14 [32] images consists of different sizes of fourteen images. BSDS100 [81] test
 388 datasets depends on 100 different natural scenes of images. Urban100 [85] is the challenging test
 389 image dataset having different frequency bands with detailed information. Manga109 [86] test image
 390 dataset depend on different comic type images with fine structures.

391 3.4. Implementation Details

392 Our proposed method has been designed, trained and tested under the Ubuntu 18.04 platform
 393 supporting with NVIDIA 25 Titan Xp GPU with 64GB RAM. We trained our model on scale
 394 enhancement factor of 2x, 4x and 8x in Keras 2.2.1 framework [87]. Quantitative evaluation like
 395 PSNR/SSIM were tested on MATLAB 2018a environment for single image super-resolution pictures.

396

9 3.5. Comparison with Other State-of-the-art-methods

397

We compare the performance of our proposed MXDSIR approach with several state-of-the-art methods, including Bicubic, SRCNN, A+, RFL, SelfExSR, SRCNN, FSRCNN, SCN, VDSR, DRCN, LapSRN, DRRN, MSISRD, MemNet.

400

Table 1 present the results of $2 \times$, $4 \times$ and $8 \times$ image super-resolution, respectively. For the $2 \times$, $4 \times$ and $8 \times$ super-resolution tasks, our proposed approach outperform as compare to other methods on most datasets. For the $8 \times$ super-resolution tasks, our proposed method achieve the best and the second best performance among all the state-of-the-art methods, respectively. These observations demonstrate the effectiveness of the proposed methods. Table 2 and 3 presents the overall average of PSNR and SSIM. Our proposed method indicates the best and second best performance on scale factor $2 \times$, $4 \times$ and $8 \times$. Only for scale factor $4 \times$ results of average PSNR of MSISRD method has little difference than our proposed method. Similarly, DRRN has little improvement in terms of SSIM on scale $4 \times$ enlargement factor.

409

Table 1. Quantitative assessment of existing deep convolutional neural network SR methods with our proposed method; results are reported as the average value of PSNR and SSIM with enlargement factor $2 \times$, $4 \times$, and $8 \times$; bold indicated result with red color is the best value, and the second best is the underlined result with blue color.

Method	Year	Factor	Params	17		BSDS100	Urban100	Manga109
				Set5	Set14			
				PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
Bicubic	1991	$2 \times$	-/-	33.69 / 0.931	30.25 / 0.870	29.57 / 0.844	26.89 / 0.841	30.86 / 0.936
A+		$2 \times$	-/-	36.60 / 0.955	32.32 / 0.906	31.24 / 0.887	29.25 / 0.895	35.37 / 0.968
RFL		$2 \times$	-/-	36.59 / 0.954	32.29 / 0.905	31.18 / 0.885	29.14 / 0.891	35.12 / 0.966
SelfExSR		$2 \times$	-/-	36.60 / 0.955	32.24 / 0.904	31.20 / 0.887	29.55 / 0.898	35.82 / 0.969
SRCNN		$2 \times$	57k	36.72 / 0.955	32.51 / 0.908	31.38 / 0.889	29.53 / 0.896	35.76 / 0.968
FSRCNN		$2 \times$	12k	37.05 / 0.956	32.66 / 0.909	31.53 / 0.892	29.88 / 0.902	36.67 / 0.971
SCN		$2 \times$	42k	36.58 / 0.954	32.35 / 0.905	31.26 / 0.885	29.52 / 0.897	35.51 / 0.967
VDSR		$2 \times$	665k	37.53 / 0.959	33.05 / 0.913	31.90 / 0.896	30.77 / 0.914	37.22 / 0.975
DRCN		$2 \times$	1775k	37.63 / 0.959	33.06 / 0.912	31.85 / 0.895	30.76 / 0.914	37.63 / 0.974
LapSRN		$2 \times$	812k	37.52 / 0.959	33.08 / 0.913	31.80 / 0.895	30.41 / 0.910	37.27 / 0.974
DRRN		$2 \times$	297k	37.74 / 0.959	33.23 / 0.914	32.05 / 0.897	31.23 / 0.919	37.92 / 0.976
MSISRD		$2 \times$	240k	37.80 / 0.960	33.84 / 0.920	32.09 / 0.895	31.10 / 0.913	37.70 / 0.975
MemNet		$2 \times$	677k	37.78 / 0.959	33.28 / 0.914	32.08 / 0.897	31.31 / 0.919	37.72 / 0.974
DPRN		$2 \times$		37.74 / 0.959				
MXDSIR		$2 \times$	210k	37.83 / 0.959	33.87 / 0.920	32.12 / 0.897	31.33 / 0.918	37.93 / 0.976
Bicubic		$4 \times$	-/-	28.43 / 0.811	26.01 / 0.704	25.97 / 0.670	23.15 / 0.660	24.93 / 0.790
A+		$4 \times$	-/-	30.32 / 0.860	27.34 / 0.751	26.83 / 0.711	24.34 / 0.721	27.03 / 0.851
RFL		$4 \times$	-/-	30.17 / 0.855	27.24 / 0.747	26.76 / 0.708	24.20 / 0.712	26.80 / 0.841
SelfExSR		$4 \times$	-/-	30.34 / 0.862	27.41 / 0.753	26.84 / 0.713	24.83 / 0.740	27.83 / 0.866
SRCNN		$4 \times$	51k	30.50 / 0.863	27.52 / 0.753	26.91 / 0.712	24.53 / 0.725	27.66 / 0.859
FSRCNN		$4 \times$	12k	30.72 / 0.866	27.61 / 0.755	26.98 / 0.715	24.62 / 0.728	27.90 / 0.861
SCN		$4 \times$	42k	30.41 / 0.863	27.39 / 0.751	26.88 / 0.711	24.52 / 0.726	27.39 / 0.857
VDSR		$4 \times$	665k	31.35 / 0.883	28.02 / 0.768	27.29 / 0.726	25.18 / 0.754	28.83 / 0.887
DRCN		$4 \times$	1775k	31.54 / 0.884	28.03 / 0.768	27.24 / 0.725	25.14 / 0.752	28.98 / 0.887
LapSRN		$4 \times$	812k	31.54 / 0.885	28.19 / 0.772	27.32 / 0.727	25.21 / 0.756	29.09 / 0.890
DRRN		$4 \times$	297k	31.68 / 0.888	28.21 / 0.772	27.38 / 0.728	25.44 / 0.764	29.46 / 0.896
MSISRD		$4 \times$	240k	31.62 / 0.886	28.51 / 0.771	27.33 / 0.727	25.42 / 0.757	31.61 / 0.891
MemNet		$4 \times$	677k	31.74 / 0.889	28.26 / 0.772	27.40 / 0.728	25.50 / 0.763	29.42 / 0.894
MXDSIR		$4 \times$	210k	31.77 / 0.888	28.63 / 0.772	27.45 / 0.728	25.54 / 0.763	30.21 / 0.895
Bicubic		$8 \times$	-/-	24.40 / 0.658	23.10 / 0.566	23.67 / 0.548	20.74 / 0.516	21.47 / 0.650
A+		$8 \times$	-/-	25.53 / 0.693	23.89 / 0.595	24.21 / 0.569	21.37 / 0.546	22.39 / 0.681
RFL		$8 \times$	-/-	25.38 / 0.679	23.79 / 0.587	24.13 / 0.563	21.27 / 0.536	22.28 / 0.669
SelfExSR		$8 \times$	-/-	25.49 / 0.703	23.92 / 0.601	24.19 / 0.568	21.81 / 0.577	22.99 / 0.719
SRCNN		$8 \times$	57k	25.33 / 0.690	23.76 / 0.591	24.13 / 0.566	21.29 / 0.544	22.46 / 0.695

		²	⁵	²	⁵	²	⁵	²	⁵
FSRCNN		8x	12k	25.60 / 0.697	24.00 / 0.599	24.31 / 0.572	21.45 / 0.550	22.72 / 0.692	
SCN		8x	42k	25.59 / 0.706	24.02 / 0.603	24.30 / 0.573	21.52 / 0.560	22.68 / 0.701	
VDSR		8x	665k	25.93 / 0.724	24.26 / 0.614	24.49 / 0.583	21.70 / 0.571	23.16 / 0.725	
DRCN		8x	1775k	25.93 / 0.723	24.25 / 0.614	24.49 / 0.582	21.71 / 0.571	23.20 / 0.724	
LapSRN		8x	812k	26.15 / 0.738	24.35 / 0.620	24.54 / 0.586	21.81 / 0.581	23.39 / 0.735	
DRRN		8x	297k	26.18 / 0.738	24.42 / 0.622	24.59 / 0.587	21.88 / 0.583	23.60 / 0.742	
MSISRD		8x	240k	26.26 / 0.737	24.38 / 0.621	24.73 / 0.586	22.53 / 0.582	23.50 / 0.738	
MemNet		8x	677k	26.16 / 0.741	24.38 / 0.619	24.58 / 0.584	21.89 / 0.582	23.56 / 0.738	
MXDSIR		8x	210k	26.21 / 0.740	24.42 / 0.621	24.77 / 0.587	22.91 / 0.582	23.63 / 0.739	

413

414

Table 2. Benchmark results of average value of PSNR on test datasets Set5, Set14, BSDS100, Urban100 and Manga109 for enlargement factor 2x, 4x, and 8x; red color with bold value indicates the best value, and the blue color with underline indicates the second best value.

415

416

Method	Factor	Set5	Set14	BSDS100	Urban100	Manga109	Average
		PSNR	PSNR	PSNR	PSNR	PSNR	PSNR
Bicubic	2x	33.69	30.25	29.57	26.89	30.86	30.25
A+	2x	36.60	32.32	31.24	29.25	35.37	32.96
RFL	2x	36.59	32.29	31.18	29.14	35.12	32.86
SelfExSR	2x	36.60	32.24	31.20	29.55	35.82	33.08
SRCNN	2x	36.72	32.51	31.38	29.53	35.76	33.18
FSRCNN	2x	37.05	32.66	31.53	29.88	36.67	33.56
SCN	2x	36.58	32.35	31.26	29.52	35.51	33.04
VDSR	2x	37.53	33.05	31.90	30.77	37.22	34.09
DRCN	2x	37.63	33.06	31.85	30.76	37.63	34.19
LapSRN	2x	37.52	33.08	31.80	30.41	37.27	34.02
DRRN	2x	37.74	33.23	32.05	31.23	37.92	34.43
MSISRD	2x	37.80	33.84	32.09	31.10	37.70	34.51
MemNet	2x	37.78	33.28	32.08	31.31	37.72	34.43
DPRN [88]	2x	37.74	33.23				
MXDSIR	2x	37.83	33.87	32.12	31.33	37.93	34.62
Bicubic	4x	28.43	26.01	25.97	23.15	24.93	25.70
A+	4x	30.32	27.34	26.83	24.34	27.03	27.17
RFL	4x	30.17	27.24	26.76	24.20	26.80	27.03
SelfExSR	4x	30.34	27.41	26.84	24.83	27.83	27.45
SRCNN	4x	30.50	27.52	26.91	24.53	27.66	27.42
FSRCNN	4x	30.72	27.61	26.98	24.62	27.90	27.57
SCN	4x	30.41	27.39	26.88	24.52	27.39	27.32
VDSR	4x	31.35	28.02	27.29	25.18	28.83	28.13
DRCN	4x	31.54	28.03	27.24	25.14	28.98	28.19
LapSRN	4x	31.54	28.19	27.32	25.21	29.09	28.27
DRRN	4x	31.68	28.21	27.38	25.44	29.46	28.43
MSISRD	4x	31.62	28.51	27.33	25.42	31.61	28.90
MemNet	4x	31.74	28.26	27.40	25.50	29.42	28.46
MXDSIR	4x	31.77	28.63	27.45	25.54	30.21	28.72
Bicubic	8x	24.40	23.10	23.67	20.74	21.47	22.68
A+	8x	25.53	23.89	24.21	21.37	22.39	23.48
RFL	8x	25.38	23.79	24.13	21.27	22.28	23.37
SelfExSR	8x	25.49	23.92	24.19	21.81	22.99	23.68
SRCNN	8x	25.33	23.76	24.13	21.29	22.46	23.39
FSRCNN	8x	25.60	24.00	24.31	21.45	22.72	23.62
SCN	8x	25.59	24.02	24.30	21.52	22.68	23.62
VDSR	8x	25.93	24.26	24.49	21.70	23.16	23.91
DRCN	8x	25.93	24.25	24.49	21.71	23.20	23.92
LapSRN	8x	26.15	24.35	24.54	21.81	23.39	24.05

DRRN	8×	26.18	24.42	24.59	21.88	23.60	24.13
MSISRD	8×	26.26	24.38	24.73	22.53	23.50	<u>24.28</u>
MemNet	8×	26.16	24.38	24.58	21.89	23.56	24.11
MXDSIR	8×	26.21	24.42	24.77	22.91	23.63	24.39

418

419
420
421

Table 3. Benchmark results of average value of SSIM on test datasets Set5, Set14, BSDS100, Urban100 and Manga109 for enlargement factor 2×, 4×, and 8×; red color with bold value indicates the best value, and the blue color with underline indicates the second best value.

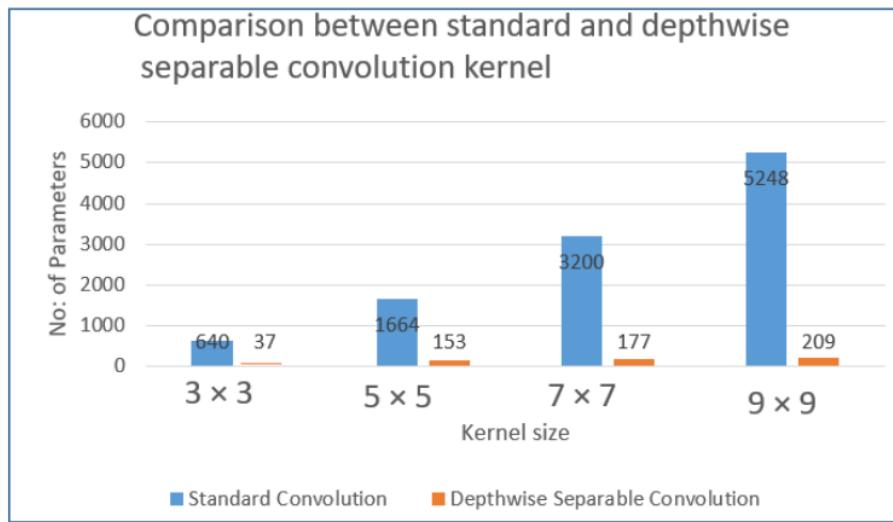
422

Method	Factor	Set5	Set14	BSDS100	Urban100	Manga109	Average
		SSIM	SSIM	SSIM	SSIM	SSIM	SSIM
Bicubic	2×	0.931	0.870	0.844	0.841	0.936	0.884
A+	2×	0.955	0.906	0.887	0.895	0.968	0.922
RFL	2×	0.954	0.905	0.885	0.891	0.966	0.920
SelfExSR	2×	0.955	0.904	0.887	0.898	0.969	0.923
SRCNN	2×	0.955	0.908	0.889	0.896	0.968	0.923
FSRCNN	2×	0.956	0.909	0.892	0.902	0.971	0.926
SCN	2×	0.954	0.905	0.885	0.897	0.967	0.922
VDSR	2×	0.959	0.913	0.896	0.914	0.975	0.931
DRCN	2×	0.959	0.912	0.895	0.914	0.974	0.931
LapSRN	2×	0.959	0.913	0.895	0.910	0.974	0.930
DRRN	2×	0.959	0.914	0.897	0.919	0.976	<u>0.933</u>
MSISRD	2×	0.960	0.920	0.895	0.913	0.975	<u>0.933</u>
MemNet	2×	0.959	0.914	0.897	0.919	0.974	<u>0.933</u>
MXDSIR	2×	0.959	0.920	0.897	0.918	0.976	0.934
Bicubic	4×	0.811	0.704	0.670	0.660	0.790	0.727
A+	4×	0.860	0.751	0.711	0.721	0.851	0.779
RFL	4×	0.855	0.747	0.708	0.712	0.841	0.773
SelfExSR	4×	0.862	0.753	0.713	0.740	0.866	0.787
SRCNN	4×	0.863	0.753	0.712	0.725	0.859	0.782
FSRCNN	4×	0.866	0.755	0.715	0.728	0.861	0.785
SCN	4×	0.863	0.751	0.711	0.726	0.857	0.782
VDSR	4×	0.883	0.768	0.726	0.754	0.887	0.804
DRCN	4×	0.884	0.768	0.725	0.752	0.887	0.803
LapSRN	4×	0.885	0.772	0.727	0.756	0.890	0.806
DRRN	4×	0.888	0.772	0.728	0.764	0.896	0.810
MSISRD	4×	0.886	0.771	0.727	0.757	0.891	0.806
MemNet	4×	0.889	0.772	0.728	0.763	0.894	<u>0.809</u>
MXDSIR	4×	0.888	0.772	0.728	0.763	0.895	<u>0.809</u>
Bicubic	8×	0.658	0.566	0.548	0.516	0.650	0.588
A+	8×	0.693	0.595	0.569	0.546	0.681	0.617
RFL	8×	0.679	0.587	0.563	0.536	0.669	0.607
SelfExSR	8×	0.703	0.601	0.568	0.577	0.719	0.634
SRCNN	8×	0.690	0.591	0.566	0.544	0.695	0.617
FSRCNN	8×	0.697	0.599	0.572	0.550	0.692	0.622
SCN	8×	0.706	0.603	0.573	0.560	0.701	0.629
VDSR	8×	0.724	0.614	0.583	0.571	0.725	0.643
DRCN	8×	0.723	0.614	0.582	0.571	0.724	0.643
LapSRN	8×	0.738	0.620	0.586	0.581	0.735	0.652
DRRN	8×	0.738	0.622	0.587	0.583	0.742	0.654
MSISRD	8×	0.737	0.621	0.586	0.582	0.738	<u>0.653</u>
MemNet	8×	0.741	0.619	0.584	0.582	0.738	<u>0.653</u>
MXDSIR	8×	0.740	0.621	0.587	0.582	0.739	0.654

423

3.6. Performance Comparison in terms of Kernel Size

424 Size and type of convolution kernel plays very important role in terms model size and
 425 computational cost. In Figure 5, we select the two different convolution kernels, one is standard
 426 convolution kernel and other is depthwise separable convolution kernel, with same 64 number of
 427 feature maps. Results clearly observed that our proposed depthwise separable convolution kernel
 428 has less computationally cost as compared to standard convolution operation, because depthwise
 429 separable convolution is the most efficient kernel, rather than the standard kernel.



430
 431 Figure 5. Complexity comparison between standard convolution versus depthwise separable
 432 convolution.
 433

3.7. Quantitative Comparisons in terms of average PSNR (dB) on five benchmark datasets

434 In Table 4, we provide a summary of quantitative evaluation on several datasets. The results of
 435 other ten methods are the same as reported at LapSRN. Our method outperforms all previous
 436 methods in these datasets. Compare with the newest previous methods, our method can improve
 437 roughly 3.03 dB, 30.7 dB, 1.48 dB, 1.17 dB, 1.24 dB, 1.00 dB, 1.25 dB, 0.53 dB, 0.48 dB, 0.47 dB, 0.24
 438 dB, 0.01 dB and 0.23 dB on average, in comparison with Bicubic, A+, RFL, SelfExSR, SRCNN,
 439 FSRCNN, SCN, VDSR, DRCN, LapSRN, DRRN, MSISRD and MemNet respectively.

440 Table 4. Benchmark results of average value of PSNR (dB) on test datasets Set5, Set14, BSDS100, Urban100
 441 and Manga109 for enlargement factor 2×, 4×, and 8×.

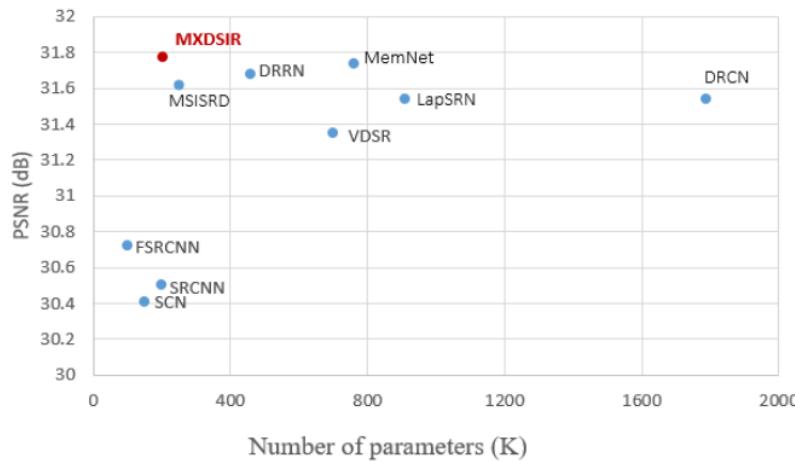
Dataset	Factor	Bicubic	A+	RFL	SelfExSR	SRCNN	FSRCNN	SCN	VDSR	DRCN	LapSRN	DRRN	MSISRD	MemNet	MXDSIR
Set5	2×	33.69	36.60	36.59	36.60	36.72	37.05	36.58	37.53	37.63	37.52	37.74	37.80	37.78	37.83
	4×	28.43	30.32	30.17	30.34	30.50	30.72	30.41	31.35	31.54	31.54	31.68	31.62	31.74	31.77
	8×	24.40	25.53	25.38	25.49	25.33	25.60	25.59	25.93	25.93	26.15	26.18	26.26	26.16	26.21
Set14	2×	30.25	32.32	32.29	32.24	32.51	32.66	32.35	33.05	33.06	33.08	33.23	33.84	33.28	33.87
	4×	26.01	27.34	27.24	27.41	27.52	27.61	27.39	28.02	28.03	28.19	28.21	28.51	28.26	28.63
BSDS100	2×	23.10	23.89	23.79	23.92	23.76	24.00	24.02	24.26	24.25	24.35	24.42	24.38	24.38	24.42
	4×	29.57	31.24	31.18	31.20	31.38	31.53	31.26	31.90	31.85	31.80	32.05	32.09	32.08	32.12
	8×	25.97	26.83	26.76	26.84	26.91	26.98	26.88	27.29	27.24	27.32	27.38	27.33	27.40	27.45
URBAN100	2×	23.67	24.21	24.13	24.19	24.13	24.31	24.30	24.49	24.49	24.54	24.59	24.73	24.58	24.77
	4×	26.89	29.25	29.14	29.55	29.53	29.88	29.52	30.77	30.76	30.41	31.23	31.10	31.31	31.33

MANGA109	8x	20.74	21.37	21.27	21.81	21.29	21.45	21.52	21.70	21.71	21.81	21.88	22.53	21.89	22.91
	2x	30.86	35.37	35.12	35.82	35.76	36.67	35.51	37.22	37.63	37.27	37.92	37.70	37.72	37.93
	4x	24.93	27.03	26.80	27.83	27.66	27.90	27.39	28.83	28.98	29.09	29.46	31.61	29.42	30.21
	8x	21.47	22.39	22.28	22.99	22.46	22.72	22.68	23.16	23.20	23.39	23.60	23.50	23.56	23.63
	Average	26.21	27.87	27.76	28.07	28.00	28.24	27.99	28.71	28.76	28.77	29.00	29.23	29.01	29.24

442

443 *3.8. Comparison in terms of Number of Model Parameters* 8

444 We show the reconstruction performance versus the number of network parameters of CNN-
 445 based SR methods as shown in Figure 8. By using depthwise separable convolution layer, our
 446 proposed model has less number of parameters as compared to state-of-the-art methods. Our
 447 MXDSIR has parameters about 68% less than the VDSR, 88% less than the DRCN, 74% less than
 448 the LapSRN, 29% less than the DRRN, 12% less than MSISR and 69% less than the MemNet.



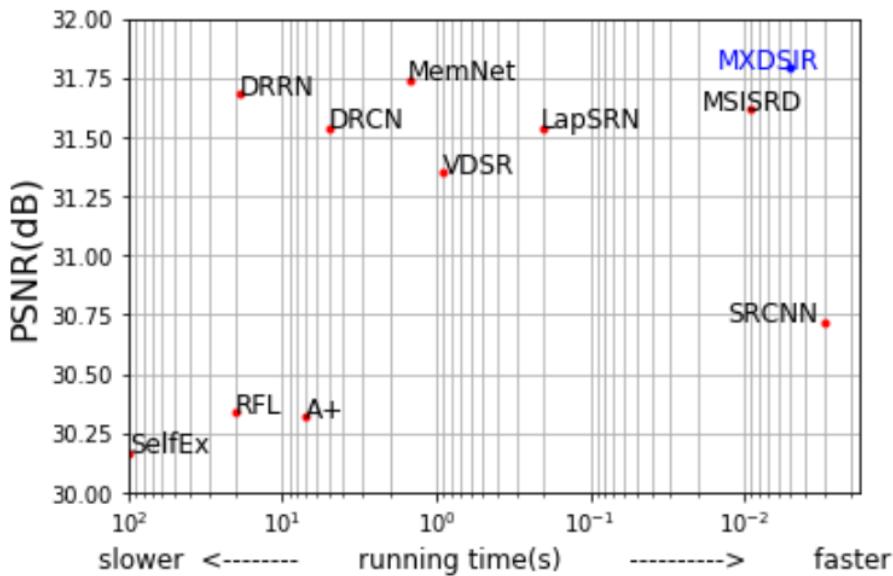
449

450 22 Figure 6. Number of network parameters versus performance. The results are evaluated on the
451 Set5 dataset for 4x SR.452 *3.9. Comparison in terms of Run time Versus PSNR* 1

453 In this section, we evaluate the performance of our proposed model in terms of running or
 454 execution time versus PSNR. The running time is calculated by the average of the Set14 for each
 455 method. As for execution time performance, we used the publicly available codes provided by the
 456 authors to compare with state-of-the-art methods on a same machine with 3.6GHz Intel i7 CPU
 457 (32GB RAM) and NVIDIA GTX 1080ti GPU (11 GB Memory). Figure.1 shows the trade-off between
 458 the execution time and performance on Set5 dataset for 4x SR. Our method is 0.09dB higher than
 459 DRRN on PSNR and approximately 10 times faster.

460

461



462

463

Figure 7. Quantitative comparison between the PSNR performance vs. runtime on Set5 scale 4x enlargement.

464

465

Figure 8 show the perceptual quality performance on the Set5 [51], Set14 [63], BSDS100 [58] and Urban100 [23] datasets for scale 4x enlargements image SR. Figure 9 present the visual performance of above datasets on scale factor 8x, including one image from the Manga109 [14] dataset. The results of the Bicubic, SRCNN [26], and FSRCNN [15] looks blurry and lack high-frequency details. Image SR on scale 8x is a very challenging problem, but our method accurately reconstructs the texture details, suppresses the artifacts, and recovers the details of the LR image with sharp edges.

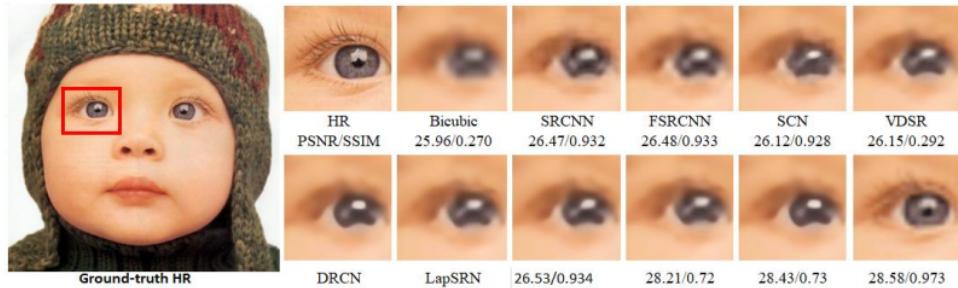
466

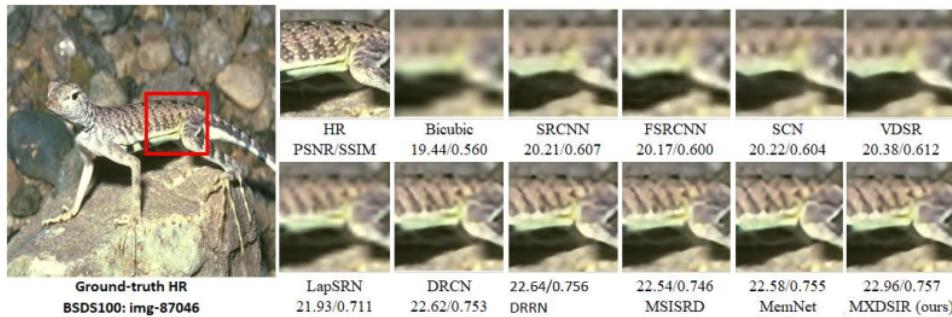
467

468

469

470





473

474 **Figure 8.** Visual performance of image on Set5, Set14, BSDS100 and Urban100 dataset with
 475 4x scale factor enlargements.

476 4. Conclusions

477 In this paper, we have presented fast and computationally efficient Xception based residual
 478 CNN for image super-resolution to extract the features directly from the original input LR input, and
 479 to reconstruct the HR image. The proposed network architecture is inspired by GoogLeNet's
 480 Xception module to produce multiple features during the feature extraction and reconstruction
 481 processes. The proposed strategies ensure that the network demonstrates fast convergence speed and
 482 low computational cost, due to replacing the interpolation upscaling with the learned transposed
 483 convolution layer. The proposed network architecture consists of an input residual skip connection
 484 layers, upsampling layer, the well-designed Xception block, and an output layer. In the features
 485 extraction part of our proposed method are sent the local as well as global features by skip connection
 486 to the reconstruction network part. Reconstruction part of the network used the Xception block,
 487 which is adopted from GoogLeNet architecture, to obtain a multiscale information with less number
 488 of parameters. We apply the high learning rate to speed up the training convergence process.
 489 Furthermore, we used depthwise separable convolution followed by parametric rectified linear unit
 490 (PReLU) activation function to increase computational efficiency and to avoid the dead features
 491 during the training process. Moreover, our network architecture is relatively simple in nature and
 492 well designed for image and computer vision tasks. Extensive experimental results on a different
 493 images are drawn from five representative image datasets, shows that the our algorithm not only
 494 provide satisfactory image SR performance quantitatively, but also has favorable results in terms of
 495 complexity and visual pleasing quality in comparison to existing state-of-the-art SR methods.
 496

497 **“Author Contributions:** Conceptualization, investigation, software, writing-original draft preparation , W.M;
 498 supervision, writing-review and editing S.A.”

499 **“Funding:** The research grant have been funded by “The 100th Anniversary Chulalongkorn University Fund
 500 for Doctoral Scholarship”.

501 **Acknowledgments:** In this section you can acknowledge any support given which is not covered by the author
 502 contribution or funding sections. This may include administrative and technical support, or donations in kind
 503 (e.g., materials used for experiments).”

504 **Conflicts of Interest:** The authors declare no conflicts of interest.”

505

506

507 **References**

- 508 1. Ren, S., et al. *Faster r-cnn: Towards real-time object detection with region proposal*
509 *networks.* in *Advances in neural information processing systems.* 2015.
- 510 2. Redmon, J., et al. *You only look once: Unified, real-time object detection.* in *Proceedings of*
511 *the IEEE conference on computer vision and pattern recognition.* 2016.
- 512 3. He, K., et al. *Mask r-cnn.* in *Proceedings of the IEEE international conference on computer*
513 *vision.* 2017.
- 514 4. Badrinarayanan, V., A. Kendall, and R. Cipolla, *Segnet: A deep convolutional encoder-decoder*
515 *architecture for image segmentation.* IEEE transactions on pattern analysis and machine
516 intelligence, 2017. **39**(12): p. 2481-2495.
- 517 5. Krizhevsky, A., I. Sutskever, and G.E. Hinton. *Imagenet classification with deep convolutional*
518 *neural networks.* in *Advances in neural information processing systems.* 2012.
- 519 6. Simonyan, K. and A. Zisserman, *Very deep convolutional networks for large-scale image*
520 *recognition.* arXiv preprint arXiv:1409.1556, 2014.
- 521 7. Tajbakhsh, N., et al., *Convolutional neural networks for medical image analysis: Full training*
522 *or fine tuning?* IEEE transactions on medical imaging, 2016. **35**(5): p. 1299-1312.
- 523 8. Peled, S. and Y. Yeshurun, *Superresolution in MRI: application to human white matter fiber*
524 *tract visualization by diffusion tensor imaging.* Magnetic Resonance in Medicine: An Official
525 Journal of the International Society for Magnetic Resonance in Medicine, 2001. **45**(1): p. 29-
526 35.
- 527 9. Shi, W., et al. *Cardiac image super-resolution with global correspondence using multi-atlas*
528 *patchmatch.* in *International Conference on Medical Image Computing and Computer-*
529 *Assisted Intervention.* 2013. Springer.
- 530 10. Schroff, F., D. Kalenichenko, and J. Philbin. *Facenet: A unified embedding for face recognition*
531 *and clustering.* in *Proceedings of the IEEE conference on computer vision and pattern*
532 *recognition.* 2015.
- 533 11. Gunturk, B.K., et al., *Eigenface-domain super-resolution for face recognition.* IEEE
534 *transactions on image processing, 2003.* **12**(5): p. 597-606.
- 535 12. Goto, T., et al. *Super-resolution System for 4K-HDTV.* in *2014 22nd International Conference*
536 *on Pattern Recognition.* 2014. IEEE.
- 537 13. Zhang, L., et al., *A super-resolution reconstruction algorithm for surveillance images.* Signal
538 *Processing, 2010.* **90**(3): p. 848-859.
- 539 14. Thornton, M.W., P.M. Atkinson, and D. Holland, *Sub-pixel mapping of rural land cover*
540 *objects from fine spatial resolution satellite sensor imagery using super-resolution pixel-*
541 *swapping.* International Journal of Remote Sensing, 2006. **27**(3): p. 473-491.
- 542 15. Uçar, A., Y. Demir, and C. Güzelış, *Object recognition and detection with deep learning for*
543 *autonomous driving applications.* Simulation, 2017. **93**(9): p. 759-769.
- 544 16. Pelliccione, P., et al., *Automotive architecture framework: The experience of volvo cars.*
545 *Journal of systems architecture, 2017.* **77**: p. 83-100.
- 546 17. Nasrollahi, K. and T.B. Moeslund, *Super-resolution: a comprehensive survey.* Machine vision
547 *and applications, 2014.* **25**(6): p. 1423-1468.
- 548 18. Duchon, C.E., *Lanczos filtering in one and two dimensions.* Journal of applied meteorology,
549 1979. **18**(8): p. 1016-1022.

- 550 19. Zhang, L. and X. Wu, *An edge-guided image interpolation algorithm via directional filtering* and data fusion. IEEE transactions on Image Processing, 2006. **15**(8): p. 2226-2238.
- 551 20. Wang, L., et al., *Edge-directed single-image super-resolution via adaptive gradient* magnitude self-interpolation. IEEE Transactions on Circuits and Systems for Video Technology, 2013. **23**(8): p. 1289-1299.
- 552 21. Sun, J., Z. Xu, and H.-Y. Shum. *Image super-resolution using gradient profile prior.* in 2008 IEEE Conference on Computer Vision and Pattern Recognition. 2008. IEEE.
- 553 22. Huang, J.-J. and W.-C. Siu, *Learning Hierarchical Decision Trees for Single-Image Super-* Resolution. IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, 2017. **27**(5): p. 937.
- 554 23. Lin, Z., et al., *Limits of learning-based superresolution algorithms.* International journal of computer vision, 2008. **80**(3): p. 406-420.
- 555 24. Freeman, W.T., E.C. Pasztor, and O.T. Carmichael, *Learning low-level vision.* International journal of computer vision, 2000. **40**(1): p. 25-47.
- 556 25. Freeman, W.T., T.R. Jones, and E.C. Pasztor, *Example-based super-resolution.* IEEE Computer graphics and Applications, 2002(2): p. 56-65.
- 557 26. Yang, C.-Y., C. Ma, and M.-H. Yang. *Single-image super-resolution: A benchmark.* in European Conference on Computer Vision. 2014. Springer.
- 558 27. Gao, X., et al., *Image super-resolution with sparse neighbor embedding.* IEEE Transactions on Image Processing, 2012. **21**(7): p. 3194-3205.
- 559 28. Mishra, D., et al., *Development of robust neighbor embedding based super-resolution scheme.* Neurocomputing, 2016. **202**: p. 49-66.
- 560 29. Chang, H., D.-Y. Yeung, and Y. Xiong. *Super-resolution through neighbor embedding.* in Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. 2004. IEEE.
- 561 30. Jianchao, Y., et al. *Image super-resolution as sparse representation of raw image patches.* in Proc. IEEE Conf. on Computer Vision and Pattern Recognition. 2008.
- 562 31. Yang, J., et al., *Image super-resolution via sparse representation.* IEEE transactions on image processing, 2010. **19**(11): p. 2861-2873.
- 563 32. Zeyde, R., M. Elad, and M. Protter. *On single image scale-up using sparse-representations.* in International conference on curves and surfaces. 2010. Springer.
- 564 33. Gibbon, K.T. and D.G. Bailey. *A novel approach to real-time bilinear interpolation.* in Proceedings. DELTA 2004. Second IEEE International Workshop on Electronic Design, Test and Applications. 2004. IEEE.
- 565 34. Keys, R., *Cubic convolution interpolation for digital image processing.* IEEE transactions on acoustics, speech, and signal processing, 1981. **29**(6): p. 1153-1160.
- 566 35. Dai, S., et al., *Softcuts: a soft edge smoothness prior for color image super-resolution.* IEEE Transactions on Image Processing, 2009. **18**(5): p. 969-981.
- 567 36. Yan, Q., et al., *Single image superresolution based on gradient profile sharpness.* IEEE Transactions on Image Processing, 2015. **24**(10): p. 3187-3202.
- 568 37. Kim, K.I. and Y. Kwon, *Example-based learning for single-image super-resolution and JPEG artifact removal.* 2008.
- 569 38. Tai, Y.-W., et al., *Super resolution using edge prior and single image detail synthesis.* 2010.

- 593 39. Aharon, M., M. Elad, and A. Bruckstein, *K-SVD: An algorithm for designing overcomplete*
- 594 *dictionaries for sparse representation*. IEEE Transactions on signal processing, 2006. **54**(11):
- 595 p. 4311-4322.
- 596 40. Timofte, R., V. De Smet, and L. Van Gool. *Anchored neighborhood regression for fast*
- 597 *example-based super-resolution*. in *Proceedings of the IEEE international conference on*
- 598 *computer vision*. 2013.
- 599 41. Timofte, R., V. De Smet, and L. Van Gool. *A+: Adjusted anchored neighborhood regression*
- 600 *for fast super-resolution*. in *Asian Conference on Computer Vision*. 2014. Springer.
- 601 42. Cao, F., et al., *Image Super-Resolution via Adaptive $\|\cdot\|_p$ Regularization and*
- 602 *Sparse Representation*. IEEE transactions on neural networks and learning systems, 2016.
- 603 **27**(7): p. 1550-1561.
- 604 43. Liu, J., et al., *Retrieval compensated group structured sparsity for image super-resolution*.
IEEE Transactions on Multimedia, 2016. **19**(2): p. 302-316.
- 605 44. Zhang, K., et al., *Coarse-to-fine learning for single-image super-resolution*. IEEE transactions
on neural networks and learning systems, 2016. **28**(5): p. 1109-1122.
- 606 45. Yu, J., et al., *A unified learning framework for single image super-resolution*. IEEE
Transactions on Neural networks and Learning systems, 2013. **25**(4): p. 780-792.
- 607 46. Deng, C., et al., *Similarity Constraints-Based Structured Output Regression Machine: An*
Approach to Image Super-Resolution. 2016.
- 608 47. Yang, W., et al., *Consistent coding scheme for single-image super-resolution via independent*
dictionaries. IEEE Transactions on Multimedia, 2016. **18**(3): p. 313-325.
- 609 48. Kim, J., J. Kwon Lee, and K. Mu Lee. *Accurate image super-resolution using very deep*
convolutional networks. in *Proceedings of the IEEE conference on computer vision and*
pattern recognition. 2016.
- 610 49. Kim, J., J. Kwon Lee, and K. Mu Lee. *Deeply-recursive convolutional network for image super-*
resolution. in *Proceedings of the IEEE conference on computer vision and pattern*
recognition. 2016.
- 611 50. Ledig, C., et al., *Photo-realistic single image super-resolution using a generative adversarial*
network. arXiv preprint, 2017.
- 612 51. Dong, C., et al., *Image super-resolution using deep convolutional networks*. IEEE transactions
on pattern analysis and machine intelligence, 2016. **38**(2): p. 295-307.
- 613 52. Dong, C., C.C. Loy, and X. Tang. *Accelerating the super-resolution convolutional neural*
network. in *European Conference on Computer Vision*. 2016. Springer.
- 614 53. Shi, W., et al. *Real-time single image and video super-resolution using an efficient sub-pixel*
convolutional neural network. in *Proceedings of the IEEE Conference on Computer Vision and*
Pattern Recognition. 2016.
- 615 54. Wang, Z., et al. *Deep networks for image super-resolution with sparse prior*. in *Proceedings*
of the IEEE International Conference on Computer Vision. 2015.
- 616 55. Wang, Z., et al., *Deeply improved sparse coding for image super-resolution*. arXiv preprint
arXiv:1507.08905, 2015. **2**(3): p. 4.
- 617 56. Mao, X., C. Shen, and Y.-B. Yang. *Image restoration using very deep convolutional encoder-*
decoder networks with symmetric skip connections. in *Advances in neural information*
processing systems. 2016.

- 636 57. Romano, Y., J. Isidoro, and P. Milanfar, *RAISR: rapid and accurate image super resolution*.
637 IEEE Transactions on Computational Imaging, 2016. **3**(1): p. 110-125.
- 638 58. Lai, W.-S., et al. *Deep laplacian pyramid networks for fast and accurate super-resolution*. in
639 *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
- 640 59. Zhang, K., et al., *Beyond a gaussian denoiser: Residual learning of deep cnn for image*
641 *denoising*. IEEE Transactions on Image Processing, 2017. **26**(7): p. 3142-3155.
- 642 60. Zhao, Y., et al., *GUN: Gradual upsampling network for single image super-resolution*. IEEE
643 Access, 2018. **6**: p. 39363-39374.
- 644 61. Tai, Y., J. Yang, and X. Liu. *Image super-resolution via deep recursive residual network*. in
645 *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017.
- 646 62. Lim, B., et al. *Enhanced deep residual networks for single image super-resolution*. in *The IEEE*
647 *conference on computer vision and pattern recognition (CVPR) workshops*. 2017.
- 648 63. Timofte, R., et al. *Ntire 2017 challenge on single image super-resolution: Methods and*
649 *results*. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*
650 *Workshops*. 2017.
- 651 64. Tai, Y., et al. *Memnet: A persistent memory network for image restoration*. in *Proceedings*
652 *of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017.
- 653 65. Yamanaka, J., S. Kuwashima, and T. Kurita. *Fast and accurate image super resolution by deep*
654 *CNN with skip connection and network in network*. in *Neural Information Processing*. 2017.
655 Springer.
- 656 66. Muhammad, W. and S. Aramvith, *Multi-Scale Inception Based Super-Resolution Using Deep*
657 *Learning Approach*. Electronics, 2019. **8**(8): p. 892.
- 658 67. Li, J., et al. *Multi-scale residual network for image super-resolution*. in *Proceedings of the*
659 *European Conference on Computer Vision (ECCV)*. 2018.
- 660 68. Wang, Y., et al., *End-to-end image super-resolution via deep and shallow convolutional*
661 *networks*. arXiv preprint arXiv:1607.07680, 2016.
- 662 69. Nair, V. and G.E. Hinton. *Rectified linear units improve restricted boltzmann machines*. in
663 *Proceedings of the 27th international conference on machine learning (ICML-10)*. 2010.
- 664 70. He, K., et al. *Delving deep into rectifiers: Surpassing human-level performance on imangenet*
665 *classification*. in *Proceedings of the IEEE international conference on computer vision*. 2015.
- 666 71. Hui, Z., X. Wang, and X. Gao. *Fast and Accurate Single Image Super-Resolution via*
667 *Information Distillation Network*. in *Proceedings of the IEEE Conference on Computer Vision*
668 *and Pattern Recognition*. 2018.
- 669 72. He, K., et al. *Deep residual learning for image recognition*. in *Proceedings of the IEEE*
670 *conference on computer vision and pattern recognition*. 2016.
- 671 73. Ioffe, S. and C. Szegedy, *Batch normalization: Accelerating deep network training by*
672 *reducing internal covariate shift*. arXiv preprint arXiv:1502.03167, 2015.
- 673 74. Szegedy, C., et al. *Rethinking the inception architecture for computer vision*. in *Proceedings*
674 *of the IEEE conference on computer vision and pattern recognition*. 2016.
- 675 75. Lin, M., Q. Chen, and S. Yan, *Network in network*. arXiv preprint arXiv:1312.4400, 2013.
- 676 76. Dong, C., et al., *Image super-resolution using deep convolutional networks*. IEEE transactions
677 on pattern analysis and machine intelligence, 2015. **38**(2): p. 295-307.

- 678 77. Chen, Y. and T. Pock, *Trainable nonlinear reaction diffusion: A flexible framework for fast*
679 *and effective image restoration*. IEEE transactions on pattern analysis and machine
680 intelligence, 2016. **39**(6): p. 1256-1272.
- 681 78. Chollet, F., *Xception: Deep learning with depthwise separable convolutions*. arXiv preprint,
682 2017: p. 1610.02357.
- 683 79. Sifre, L. and S. Mallat, *Rigid-motion scattering for image classification*. Ph. D. dissertation,
684 2014.
- 685 80. Howard, A.G., et al., *Mobilenets: Efficient convolutional neural networks for mobile vision*
686 *applications*. arXiv preprint arXiv:1704.04861, 2017.
- 687 81. Arbelaez, P., et al., *Contour detection and hierarchical image segmentation*. IEEE
688 transactions on pattern analysis and machine intelligence, 2010. **33**(5): p. 898-916.
- 689 82. Tai, Y., J. Yang, and X. Liu. *Image super-resolution via deep recursive residual network*. in
690 *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
- 691 83. Kingma, D. and L. Ba, *Adam: A Method for Stochastic Optimization*. 2015.
- 692 84. Bevilacqua, M., et al., *Low-complexity single-image super-resolution based on nonnegative*
693 *neighbor embedding*. 2012.
- 694 85. Huang, J.-B., A. Singh, and N. Ahuja. *Single image super-resolution from transformed self-*
695 *exemplars*. in *Proceedings of the IEEE Conference on Computer Vision and Pattern*
696 *Recognition*. 2015.
- 697 86. Matsui, Y., et al., *Sketch-based manga retrieval using manga109 dataset*. 2016.
- 698 87. Chollet, F., *Keras documentation*. keras. io, 2015.
- 699 88. Sha, F., S.M. Zandavi, and Y.Y. Chung, *Fast deep parallel residual network for accurate super*
700 *resolution image processing*. Expert Systems with Applications, 2019. **128**: p. 157-168.

701



© 2019 by the authors. Submitted for possible open access publication under the terms
and conditions of the Creative Commons Attribution (CC BY) license
(<http://creativecommons.org/licenses/by/4.0/>).

702

Xception

ORIGINALITY REPORT



PRIMARY SOURCES

- | | | |
|---|--|-----|
| 1 | Wazir Muhammad, Supavadee Aramvith, Takao Onoye. "Multi-scale Xception based depthwise separable convolution for single image super-resolution", PLOS ONE, 2021
Publication | 26% |
| 2 | www.mdpi.com
Internet Source | 5% |
| 3 | Submitted to University of Sheffield
Student Paper | 4% |
| 4 | Submitted to Democritus University
Student Paper | 2% |
| 5 | export.arxiv.org
Internet Source | 1 % |
| 6 | Wuzhen Shi, Feng Jiang, Debin Zhao. "Single image super-resolution with dilated convolution based multi-scale information learning inception module", 2017 IEEE International Conference on Image Processing (ICIP), 2017
Publication | 1 % |

- 7 Wazir Muhammad, Zuhaiabuddin Bhutto, Arslan Ansari, Mudasar Latif Memon et al. "Multi-Path Deep CNN with Residual Inception Network for Single Image Super-Resolution", *Electronics*, 2021 1 %
Publication
-
- 8 escholarship.org <1 %
Internet Source
-
- 9 "Computer Vision – ECCV 2018 Workshops", Springer Science and Business Media LLC, 2019 <1 %
Publication
-
- 10 Wei Ding, Zeyu Huang, Zunkai Huang, Li Tian, Hui Wang, Songlin Feng. "Designing efficient accelerator of depthwise separable convolutional neural network on FPGA", *Journal of Systems Architecture*, 2019 <1 %
Publication
-
- 11 "Neural Information Processing", Springer Science and Business Media LLC, 2017 <1 %
Publication
-
- 12 Chao Dong, Chen Change Loy, Xiaoou Tang. "Chapter 25 Accelerating the Super-Resolution Convolutional Neural Network", Springer Science and Business Media LLC, 2016 <1 %
Publication
-
- 13 openaccess.thecvf.com

<1 %

14 [deepai.org](#) <1 %
Internet Source

15 [orca.cf.ac.uk](#) <1 %
Internet Source

16 [Yongliang Tang, Weiguo Gong, Xi Chen, Weihong Li. "Deep Inception-Residual Laplacian Pyramid Networks for Accurate Single-Image Super-Resolution", IEEE Transactions on Neural Networks and Learning Systems, 2020](#) <1 %
Publication

17 [Submitted to Indian Institute of Technology, Madras](#) <1 %
Student Paper

18 [Ying Tai, Jian Yang, Xiaoming Liu, Chunyan Xu. "MemNet: A Persistent Memory Network for Image Restoration", 2017 IEEE International Conference on Computer Vision \(ICCV\), 2017](#) <1 %
Publication

19 [Submitted to University of Strathclyde](#) <1 %
Student Paper

20 [Feng Sha, Seid Miad Zandavi, Yuk Ying Chung. "Fast deep parallel residual network for](#) <1 %

accurate super resolution image processing",
Expert Systems with Applications, 2019

Publication

-
- 21 link.springer.com <1 %
Internet Source
- 22 Wei-Sheng Lai, Jia-Bin Huang, Narendra Ahuja, Ming-Hsuan Yang. "Fast and Accurate Image Super-Resolution with Deep Laplacian Pyramid Networks", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019 <1 %
Publication
- 23 Yanting Hu, Jie Li, Yuanfei Huang, Xinbo Gao. "Channel-wise and Spatial Feature Modulation Network for Single Image Super-Resolution", IEEE Transactions on Circuits and Systems for Video Technology, 2019 <1 %
Publication
- 24 Submitted to University of Florida <1 %
Student Paper
- 25 "Neural Information Processing", Springer Science and Business Media LLC, 2018 <1 %
Publication
- 26 Jun Shi, Qingping Liu, Chaofeng Wang, Qi Zhang, Shihui Ying, Haoyu Xu. "Super-resolution reconstruction of MR image with a <1 %

novel residual learning network algorithm",
Physics in Medicine & Biology, 2018

Publication

-
- 27 "Advances in Multimedia Information Processing – PCM 2018", Springer Science and Business Media LLC, 2018 **<1 %**
- Publication
-
- 28 Xin Yang, Tangxin Xie, Yingqing Guo, Dake Zhou. "Remote sensing image super - resolution based on convolutional blind denoising adaptive dense connection", IET Image Processing, 2021 **<1 %**
- Publication
-
- 29 www.groundai.com **<1 %**
- Internet Source
-
- 30 Submitted to Bilkent University **<1 %**
- Student Paper
-
- 31 Feilong Cao, Huan Liu. "Single image super-resolution via multi-scale residual channel attention network", Neurocomputing, 2019 **<1 %**
- Publication
-
- 32 arxiv.org **<1 %**
- Internet Source
-
- 33 cvlab.cse.msu.edu **<1 %**
- Internet Source
-
- 34 pdfs.semanticscholar.org **<1 %**
- Internet Source

-
- 35 "Image and Graphics", Springer Science and Business Media LLC, 2017 **<1 %**
Publication
-
- 36 Furui Bai, Wen Lu, Yuanfei Huang, Lin Zha, Jiachen Yang. "Densely convolutional attention network for image super-resolution", Neurocomputing, 2019 **<1 %**
Publication
-
- 37 Xiaofang Li, Yirui Wu, Wen Zhang, Ruichao Wang, Feng Hou. "Deep learning methods in real-time image super-resolution: a survey", Journal of Real-Time Image Processing, 2019 **<1 %**
Publication
-
- 38 etheses.whiterose.ac.uk **<1 %**
Internet Source
-
- 39 "PRICAI 2019: Trends in Artificial Intelligence", Springer Science and Business Media LLC, 2019 **<1 %**
Publication
-
- 40 "Pattern Recognition and Computer Vision", Springer Science and Business Media LLC, 2020 **<1 %**
Publication
-
- 41 "Pattern Recognition", Springer Science and Business Media LLC, 2016 **<1 %**
Publication
-

- 42 Feng Li, Huihui Bai, Yao Zhao. "FilterNet: Adaptive Information Filtering Network for Accurate and Fast Image Super-Resolution", IEEE Transactions on Circuits and Systems for Video Technology, 2019 <1 %
Publication
-
- 43 Jingwei Xin, Nannan Wang, Xinrui Jiang, Jie Li, Heng Huang, Xinbo Gao. "Chapter 6 Binarized Neural Network for Single Image Super Resolution", Springer Science and Business Media LLC, 2020 <1 %
Publication
-
- 44 Muhammad Haris, M. Rahmat Widyanto, Hajime Nobuhara. "Inception learning super-resolution", Applied Optics, 2017 <1 %
Publication
-
- 45 Submitted to Stefan cel Mare University of Suceava <1 %
Student Paper
-
- 46 Xiaole Zhao, Yulun Zhang, Tao Zhang, Xueming Zou. "Channel Splitting Network for Single MR Image Super-Resolution", IEEE Transactions on Image Processing, 2019 <1 %
Publication
-
- 47 digitalcommons.odu.edu <1 %
Internet Source
-
- 48 researchportal.port.ac.uk

Internet Source

<1 %

49

www.researchgate.net

Internet Source

<1 %

50

"Computer Vision – ECCV 2016", Springer
Science and Business Media LLC, 2016

Publication

<1 %

51

strathprints.strath.ac.uk

Internet Source

<1 %

52

"Computer Vision – ECCV 2018", Springer
Science and Business Media LLC, 2018

Publication

<1 %

53

"Neural Information Processing", Springer
Science and Business Media LLC, 2019

Publication

<1 %

54

Lecture Notes in Computer Science, 2015.

Publication

<1 %

55

Lecture Notes in Computer Science, 2016.

Publication

<1 %

Exclude quotes

On

Exclude matches

Off

Exclude bibliography

On