IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 19, 2018, accepted September 30, 2018, date of publication October 8, 2018, date of current version November 8, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2874442

Compnet: A New Scheme for Single Image Super Resolution Based on Deep Convolutional Neural Network

ALIREZA ESMAEILZEHI, (Student Member, IEEE), M. OMAIR AHMAD, (Fellow, IEEE), AND M.N.S. SWAMY, (Fellow, IEEE)

Department of Electrical and Computer Engineering, Concordia University, Montreal, QC H3G 1M8, Canada

Corresponding author: M. Omair Ahmad (omair@ece.concordia.ca)

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada and in part by the Regroupment Strategique en Microelectronique du Quebec.

ABSTRACT The features produced by the layers of a neural network become increasingly more sparse as the network gets deeper and consequently, the learning capability of the network is not further enhanced as the number of layers is increased. In this paper, a novel residual deep network, called CompNet, is proposed for the single image super resolution problem without an excessive increase in the network complexity. The idea behind the proposed network is to compose the residual signal that is more representative of the features produced by the different layers of the network and it is not as sparse. The proposed network is experimented on different benchmark datasets and is shown to outperform the state-of-the-art schemes designed to solve the super resolution problem.

INDEX TERMS Image super resolution, residual learning, deep learning.

I. INTRODUCTION

Single image super resolution, which tries to recover the information lost because of some degradation process, is one of the challenging problems due to its ill-posed nature. Generally, the degradation process in images caused by blurring (due to the camera lens) and downsampling (due to the finite number of CCD sensors) could be modeled as a decimation process. Thus, a lossless reconstruction of the ground truth image from its degraded version is not possible. A wide range of techniques have been used to solve a single image super resolution problem. Some of the methods utilize *a priori* knowledge of the specific characteristics in images and constrain the solution of the ill-posed problem leading to a unique solution. For instance, the assumption that the high frequency content of generic images is limited [1] could be used to achieve a unique solution to the problem. Another assumption to achieve a unique solution to this problem is the availability either of the low resolution dictionary or of the high resolution dictionary [2] to optimize the other one. However, these methods try to solve this ill-posed problem, which is essentially nonlinear, by using linear techniques. There are methods such as in [3], that have used nonlinear filters to reconstruct a high resolution image from its low resolution version. Although this mapping from low to high resolution spaces is nonlinear, the methods do not provide the best results, since the mapping is not end-to-end. On the other hand, the neural network schemes have the capability of providing an end-to-end mapping.

A neural network by stacking more layers, each followed by a nonlinear activation function, creates a deep network, which increases the learning capability of the network and consequently improves its performance. In view of this and the availability of adequate computational resources, deep learning techniques have become very attractive in computer vision and various other fields. Convolutional neural nets [4], [5], which are very simple to implement, form a specific category in deep neural networks that have been demonstrated to provide very promising results [6]–[13]. In fact, deep convolutional neural nets try to extract the most important features to minimize the loss between the estimated signal and the ground truth, and therefore, provide exceptional performance.

The super resolution convolutional neural network (SRCNN) proposed in [14] is one of the pioneering works in deep neural networks. SRCNN extracts feature vectors of the bicubic interpolated low resolution image and maps them

onto a high resolution space. Then, it reconstructs a high resolution image from the feature vectors in the high resolution space. SRCNN outperforms the sparse representation-based super resolution [2]; however, it has some deficiencies. It requires considerable number of back-propagations for training a shallow network. Since the network is shallow (generally a three or four layer network), it is difficult to learn the complex structures such as delicate edges. Furthermore, the resulting high resolution image suffers from ringing effect around the edges. In addition, if the nonlinear mapping is carried out by having more layers, the final feature vectors suffer from increased sparsity that hinders in providing an appropriate reconstruction.

The fast super resolution convolutional neural net (FSRCNN) of [15] is a modified version of SRCNN, in which after extracting features from the low resolution image, the dimension of the low resolution feature vectors is reduced and the new feature vectors are transformed onto a high resolution space. Consequently, the dimension of the high resolution feature vectors is increased to that of the low resolution feature vectors. Finally, reconstruction and upsampling are carried out through a deconvolutional layer. Since FSRCNN employs several trainable interpolation kernels instead of just one fixed bicubic kernel, the performance is improved over that of SRCNN. Although FSRCNN is faster than SRCNN, it is still shallow and cannot learn complicated structures.

The sparse coding network (SCN) of [16] has been introduced to implement the sparse representation-based super resolution scheme of [2] via neural nets. It first extracts the features of a bicubic interpolated low resolution image with a convolutional layer. Then, the low resolution feature vectors are fed into a learned iterative shrinkage and thresholding algorithm (LISTA) [17] to yield sparse vectors. Next, the sparse vectors thus obtained are multiplied by the patches in a high resolution dictionary through a linear layer to form high resolution feature vectors. Finally, the estimated high resolution image is reconstructed from the high resolution feature vectors. It should be noted that SCN has the same steps as that of SRCNN, that is, feature extraction, nonlinear mapping and reconstruction, but with a different implementation. SCN contains only five layers and is still too shallow for learning the nonlinear features of the high resolution images. A modified version of SCN has been proposed in [18] and is referred to as mixture SCN (MSCN) that employs some SCNs parallel to one another for estimating the various high resolution images and provides an enhanced performance. The complexity of MSCN is comparable to that of SCN, and yet the depth of the network is not changed.

The very deep super resolution (VDSR) of [19] is one of the first studies that employs the idea of residual learning in deep networks for the image super resolution problem. VDSR uses the network to learn a residue between the ground truth and bicubic interpolated image and since these two images are correlated, the residual image has less amount of details. This property facilitates the learning process of VDSR. VDSR imitates the visual geometry group network (VGG19) of [20] in terms of the network depth. As the depth of VDSR is increased, the transformation that is carried out on the features becomes more nonlinear, and hence, the performance of the network is improved. However, similar to other deep networks, if the depth in VDSR is increased to a very large value, the residual image suffers from an increased sparsity and the performance is not further enhanced. Nevertheless, VDSR outperforms all the aforementioned networks for image super resolution in terms of both the objective and subjective metrics in view of its superior learning capability due to the use of residual image and its relatively larger depth.

Deep networks do not progressively improve the performance as the network is made more deep, since after a certain number of layers the processed results become so sparse as not to include much useful information about the image. It has been reported in [30] that: "As CNNs become increasingly deep, a new research problem emerges: as information about the input or gradient passes through many layers, it can vanish and wash out by the time it reaches the end (or beginning) of the network". ResNet [6], or its extended version [7], is a scheme that could facilitate the training process of very deep networks by controlling the aforementioned sparsity. It does so by modifying the network as a cascade of residual blocks. These residual blocks reduce the sparsity of the deeper layers that are most adversely affected by the sparsity. A network referred to as light-weight residual network for the super resolution problem has been proposed in [21] by making the network deeper and by taking care of the sparsity of the residual image using the residual blocks of ResNet in an effort to improve the performance that of VDSR. It has been shown in [6] that residual blocks prove to be quite effective for very deep networks. Thus, in order to reduce the complexity of the deep network, in [21] the widths of the layers are increased only progressively from a small value of 16 for the initial layers to finally a value of 256 for the latter ones. Although this latter effort in [21] has been made to reduce the complexity of the network, its complexity still is considerably too high to outperform VDSR.

Thus, it is clear from the above review that deep networks facilitate the learning processes in view of their ability to introduce more nonlinearity. However, as networks are made more deep, the output of the deeper layers become progressively more sparse. This is counterproductive in that it leads to increased sparsity in the deeper layers and thus hinders the learning ability of the network. On the other hand, use of residual blocks in very deep networks such as the one in [21], effectively controls the sparsity problem of very deep networks, but this control of sparsity is achieved at the expense of a substantially higher complexity of these networks. In view of this problem, it is imperative to develop techniques that enhance the learning process of the networks without making them very deep. In this paper, a new single image super resolution scheme is developed through a mechanism of residual learning of a deep convolutional neural network. The main idea of the proposed scheme is to diminish the effect of the progressively increasing sparsity in the outputs of the deeper

layers, while still benefitting from the enhanced nonlinearity of the deep network in obtaining the residual image. The main contributions of this paper can be summarized as follows:

- A new scheme for the single image super resolution problem is proposed by introducing a deep neural network architecture, CompNet, to effectively facilitate the learning process of the residual signal.
- A new scheme is devised to compose a rich residual signal by concatenating the feature maps selected from a suitably selected layers of the network.
- The design of CompNet has aimed at keeping the complexity of the deep network modest.
- Extensive experiments are carried out to demonstrate the effectiveness and superiority of the proposed CompNet over the state-of-the-arts of comparable complexity on three benchmark datasets.

The paper is organized as follows: Section II is dedicated to an overview of the works related to this study. In this section, three approaches of super resolution, namely, image sparse representation-based approach, shallow network approach and deep network approach, are briefly reviewed. In Section III, the scheme of the proposed CompNet for super resolution is introduced. The motivation and theoretical justifications for the network topology is described. The scheme used for the backpropagation and training process is also discussed. Section IV provides the experimental results of CompNet. The results on the effect of the different hyper-parameters, such as width, depth, spatial support and activation function, on the performance of the network is also investigated in this section. Finally, some conclusions of this study is provided in Section V.

II. RELATED WORKS

In this section, a classical and some of the deep learning-based image super resolution studies are briefly reviewed. These schemes yield some promising results when applied to various benchmark datasets on super resolution with different upscaling factors.

A. SPARSE REPRESENTATION BASED SUPER RESOLUTION

Sparse representation theory [22] has been widely utilized in computer vision, from classification [23], [24] to regression problems [2], [25]. In [2] the application of sparse representation in single image super resolution is proposed. According to sparse representation based super resolution, the raw bicubic interpolated low resolution patch, $y \in I\!R^m$, could be represented as a linear combination of some samples in an overcomplete low resolution dictionary, and therefore, the high resolution patch, $x \in I\!R^m$, is reconstructed from the high resolution version of these samples. The transformation from low to high resolution space is carried out with the help of a sparse vector $\boldsymbol{\alpha} \in I\!R^n$ as follows:

$$x = D_x\boldsymbol{\alpha} \quad \text{such that } \boldsymbol{\alpha} = \underset{s}{\operatorname{argmin}}(\|y - D_y\boldsymbol{s}\|^2 + \gamma\|\boldsymbol{s}\|_1) \quad (1)$$

where D_y and D_x are, respectively the low and high resolution dictionaries, and γ is a regularization parameter.

B. SUPER RESOLUTION USING SHALLOW CONVOLUTIONAL NEURAL NETWORKS

In [14], a scheme for single image super resolution has been developed using convolutional layers. A low resolution feature vector $y \in I\!R^{m_1}$ is extracted from a bicubic interpolated low resolution image by the first convolutional layer of the network. Next, this feature vector is transformed into a high resolution feature vector $x \in I\!R^{m_2}$ through another one or two convolutional layers. Since each of these convolutional layers is followed by a ReLU [12], [13], the mapping from low to high resolution space is considered as a nonlinear mapping. Generally, it is assumed that the dimension of the high resolution feature vectors is lower than that of the corresponding low resolution feature vectors, i.e. $m_1 > m_2$. Finally, the high resolution image is constructed from the high resolution feature vectors by the last convolutional layer. Thus, the architecture of the network is fully convolutional, the scheme can be applied on an image of arbitrary size.

In [16], the relation given by (1) has been implemented via linear and convolutional layers and the new scheme is referred to as sparse coding network(SCN). However, SCN utilizes the feature vectors of low resolution patches instead of their pixel intensities and yields the feature vectors of high resolution patches. Finally, the high resolution image is obtained from the high resolution feature vectors.

C. DEEP NEURAL NETWORKS FOR IMAGE SUPER RESOLUTION

In this subsection, three recent super resolution schemes based on deep neural networks with a large number of layers are reviewed. These schemes employ various strategies in designing their structures. The very deep super resolution (VDSR [19]) utilizes a deep network to predict a residual image. The deep edge guided recurrent residual learning (DEGREE) [26] employs the edges of the input image to appropriately reconstruct the high frequency content in a high resolution image. The end-to-end deep and shallow network (EEDS) [28] utilizes a combination of deep and shallow networks in order to reduce the shortcomings of deep networks.

In [19], a very deep neural network referred as to VDSR, which significantly outperforms SRCNN [14] and SCN [16], has been introduced for the super resolution problem. The output of this deep network is the residual of the ground truth and a bicubic interpolated version of the low resolution image. The residual signal, $r = X - Y$, X and Y being the high resolution and the bicubic upscaled version of the low resolution image, respectively. The main advantage of this residual deep network is that despite the residual image being sparse, it effectively enhances the learning process of the network.

Another work that uses deep networks for image super resolution is the deep edge guided recurrent residual learning (DEGREE) [26] that is based on residual learning. In DEGREE, first a low resolution image (Y) and its edges (E_y) are fed to a recurrent network. Next, the output

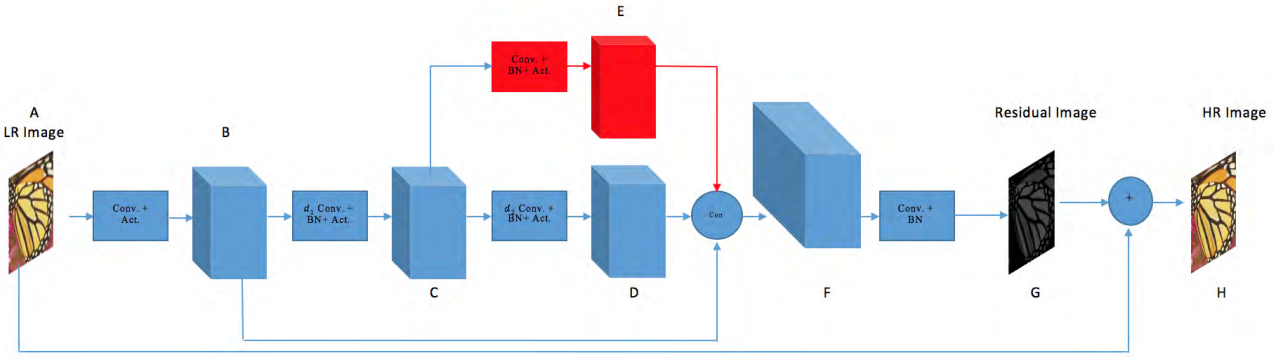


FIGURE 1. CompNet architecture. The symbols Conv., BN and Act. represent the convolution, batch normalization and activation operations, respectively. The interpolated image at node A is fed to the network. Between the nodes B and C, d_1 convolutional layers are placed. This is followed by another set of d_2 convolutional layers between the nodes C and D and a single convolutional layer between the nodes C and E. The feature maps produced at the nodes B, E and D are then concatenated through a block represented by *Con*. Feature maps at node F are fed to the last convolutional layer placed between the nodes F and G. Finally, the interpolated image from node A and the result from the node G are added to produce the estimated high resolution image.

of the recurrent net (P) is passed on to a convolutional layer to yield the edges of the high resolution image (E_x). Finally, the concatenation of P and E_x is inputted to another convolutional layer to obtain a residual (r) of high and low resolution images. Although this network provides good results, it does not outperform VDSR for some datasets like for BSD100 [27].

In [28], a combination of deep and shallow networks have been used to design an end-to-end deep and shallow (EEDS) network for image super resolution. In this study, it has been shown that the shallow network can recover the illumination lost by the deep part of network, while deep network retrieves its high frequency components. As in DEGREE [26], although EEDS outperforms SRCNN and SCN considerably, it does not supersede the performance of VDSR.

III. COMPNET FOR SINGLE IMAGE SUPER RESOLUTION

In this section, our proposed CompNet for image super resolution is developed. Here, the design of a deep network to diminish the sparsity of the deeper layers for enhancing the learning the residue between the ground truth and bicubic interpolated version of the low resolution images is carried out.

A. COMPNET: ARCHITECTURE

Learning the residual signal is the goal of the networks that are based on residual learning. The residual signal is the difference between the original high resolution image and the interpolated version of the low resolution image. The proposed deep network, referred to as CompNet, consists of several convolutional layers each followed by a ReLU activation function. The architecture of the network is shown in Fig. 1. The interpolated low resolution input signal at node A is fed to the first convolutional layer of this network, which produces at node B the first set of feature maps of the interpolated low resolution image. Between the nodes B and C, a total of d_1 convolutional layers are placed. This is followed by placing another d_2 convolutional layers between the nodes C and D and a single convolutional layer between the nodes C and E. The feature maps produced at the nodes B, E and D are then concatenated through a block represented by *Con*. Thus, node F represents a set of concatenated feature maps, which is fed to the last convolutional layer placed between the nodes F and G. Finally, the interpolated image from node A and the result from the node G are added to produce the final estimated high resolution image. The signal at node G, therefore, represents the estimated residual image. All of the layers of the network produce different feature maps of the interpolated image.

Let **u**, **v** and **w** be the feature vectors at nodes B, C and D, respectively, each produced by the convolutional layers and activation functions, say ReLU. As one progresses deeper into the network, the features produced become more sparse and they are the results of the network undergoing increasingly more nonlinearity. Thus, in the feature vectors, **u** are the least sparse and they have been produced by the network at a node where it has undergone the least nonlinearity. On the other hand, the converse is true for the feature vectors **w**.

However, depending on the value of d_1, the feature vectors represented by v have these two characteristics in between that of u and w. Thus, a feature vector that is composed by using these three feature vectors can be expected to be a better representative of the spectrum of an estimated residual image. In view of this expectation, in this investigation, the three types of vectors are concatenated as

$$c = \begin{bmatrix} w \\ v \\ u \end{bmatrix} \tag{2}$$

Let m_1, m_2' and m_3 be the numbers of slices to produce the tensors at nodes B, C and D, respectively, i.e. $u \in IR^{m_1}$, $v \in IR^{m_2'}$ and $w \in IR^{m_3}$. Then, $c \in IR^{m_1+m_2'+m_3}$. Since the sparsity of v is in between that of u and w, and amount

of the nonlinearity used to produce the feature vectors v is also in between that applied to produce u and w, we propose a dimensionality reduction of v from m_2' to m_2 by placing another convolutional layer between nodes C and E before carrying out the concatenation operation. The function of the last convolutional layer is selection of the features from u, v and w in constructing the estimated residual image.

The final feature vector, c in CompNet is expected to be less sparse in comparison to that provided by VDSR [19]. However, the feature vectors c experiences about the same amount of nonlinearity as that experienced by VDSR.

In the proposed network, a batch normalization (BN) is performed between the convolution operations and activations in each layer in order to normalize the distribution of the features produced by the activation operation of the previous layer. The batch normalization is performed by using the technique of [31] as

$$\begin{aligned} \hat{x}_i &= \frac{x_i - \mu}{\sigma}, \quad i = 1, \ldots, N_b \\ y_i &= \gamma_i \hat{x}_i + \beta_i \end{aligned} \tag{3}$$

where x_i is an element of a tensor corresponding to the *ith* sample in the batch, μ and σ are the mean and standard deviation corresponding to this element for all the samples of the batch of size N_b, γ_i and β_i are trainable parameters of BN corresponding to the *ith* sub image in the batch. Since, a batch normalization is supposed to reduce the internal covariance shift, that is produced by its previous activation function, the first layer is not required to have a batch normalization operation.

B. COMPNET: HYPERPARAMETERS

Each of the layers of the segment of the network shown in blue color between the nodes B and D uses 64 filters, each of support size 3×3. Since the purpose of the layer in the red segment of the network is dimensionality reduction, this layer employs only 32 filters each of spatial support size of 1×1. Also, since our objective is to include in the concatenation process the feature vectors whose characteristics of sparsity and nonlinearity are in between of those at nodes B and D, we chose $d_1 = d_2 = d$ and we set $d = 9$ so that network is sufficiently deep. The final convolutional layer employs a single filter of size $3 \times 3 \times 160$, for the reconstruction of the residual image. Each of the convolutional layers except the last one are followed by a ReLU activation function.

C. COMPNET: TRAINING

As in other super resolution schemes that are based on deep learning, in our scheme also sub-images are used for the training of CompNet. Sub-images of size 48×48 with no overlap are used for the training. However, since CompNet is a fully convolutional network, it can be trained and tested on images of any size. The input image is normalized to assume values in the range [0, 1]. Also the images are transformed onto the YCbCr color space. Since human eyes are more sensitive to the illumination information, the luminance channel (Y) is separated and only this channel is inputted to the network through the first layer of CompNet. The chrominance channels, Cb and Cr, are bicubic interpolated and then merged with the estimated output to construct a high resolution estimation in the YCbCr color space. Since data augmentation [33] is one of the very important processes in machine learning for boosting the performance of a network, flipped and rotated versions of the training examples are utilized to enrich the training dataset. In addition, multi-scale training is utilized for CompNet, and therefore, the training dataset consists of samples upscaled by various factors. This process not only removes the need of individual networks for each upscaling factor, but as has been shown in [19], it also improves the robustness of the network that leading to a better performance.

The effective receptive field in CompNet, when the number of layers has a default value of 20 and the filter spatial support of 3×3, is 41, that it is a little less than the size of the input sub-image. We have noticed that increasing the depth of CompNet over its default value of 20 does not improve its performance even though the corresponding effective receptive field remains within the size of the sub-images. Therefore, for the proposed CompNet, we keep the depth at 20 layers. For initializing the weights of our network, the method due to He *et al.* [34], which is based on the layer hyper parameters and the use of the ReLU activation function, is used. In this method, the kernel with a spatial support of $s \times s$ is randomly initialized with a Gaussian distribution having a zero mean and a variance of $\frac{2}{s^2 n}$, where n is the layer width.

Since the main objective metric for evaluating the single image super resolution is peak signal-to-noise ratio ($PSNR = 10 \log_{10}(\frac{255^2}{MSE})$, where MSE represents the mean squared error), the loss function used by CompNet is the mean squared error given by

$$L(\Theta) = \frac{1}{N_b} \sum_{j=1}^{N_b} \| Comp(y_j, \Theta) - x_j \|^2 + \zeta \| \Theta \|^2 \tag{4}$$

where x_j and y_j are, respectively, the *jth* ground truth and low resolution sub-images in the batch of size N_b, ζ is a weight decay parameter, Θ is the set of parameters in CompNet and $Comp(y_i, \Theta)$ is the estimate of the high resolution sub-image. For updating the parameters Θ, the stochastic gradient descent (SGD) algorithm using the Nestrov acceleration scheme [35] is adopted. The parameters of a layer are updated as

$$\begin{aligned} \theta(t) &= \theta(t-1) - \eta m(t) \\ m(t) &= \tau m(t-1) + g(t) \\ g(t) &= \nabla_{\theta(t-1)} L(\theta(t-1) - \eta \mu m(t-1) \end{aligned} \tag{5}$$

where τ and η are the momentum parameter and learning rate, respectively. In our study, the momentum parameter and initial learning rate are set as 0.9 and 0.1, respectively and the learning rate is deceased by factor of 0.1 after every 20 epochs. The skip connections that are used for feature

TABLE 1. PSNR (SSIM) Values Resulting from Applying CompNet and Various State-of-the-art Methods to Images of Three Datasets.

Dataset	Scaling	Bicubic	A+ [43]	RFL [44]	SRCNN	CSCN	VDSR	DEGREE	CompNet
Set5	×2	33.66 (0.9299)	36.54 (0.9544)	36.54(0.9537)	36.66 (0.9542)	37.00 (0.9557)	37.53 (0.9587)	37.54(0.9584)	**37.58** (**0.9596**)
	×3	30.39(0.8682)	32.58(0.9088)	32.43(0.9057)	32.75(0.9090)	33.18 (0.9153)	33.66(0.9213)	**33.72** (0.9204)	33.67(**0.9219**)
	×4	28.42(0.8104)	30.28(0.8603)	30.14(0.8548)	30.48(0.8628)	30.94(0.8755)	31.35(**0.8838**)	**31.43** (0.8818)	31.35 (0.8833)
Set14	×2	30.24(0.8688)	32.28(0.9056)	32.26(0.9040)	32.42(0.9063)	32.65 (0.9081)	33.03(0.9124)	33.01 (0.9118)	**33.29** (**0.9149**)
	×3	27.21(0.7385)	29.13(0.8188)	29.05(0.8164)	29.28(0.8209)	29.41 (0.8234)	29.77(0.8314)	29.87(0.8317)	**30.06** (**0.8368**)
	×4	26.00(0.7027)	27.32(0.7491)	27.24(0.7451)	27.49(0.7503)	27.71 (0.7592)	28.01(0.7674)	28.02 (0.7646)	**28.26** (**0.7732**)
BSD100	×2	29.56(0.8431)	31.21(0.8863)	31.16(0.8840)	31.36(0.8879)	31.46 (0.8891)	31.90(0.8960)	31.76 (0.8939)	**31.91** (**0.8972**)
	×3	27.21(0.7385)	28.29(0.7835)	28.22(0.7806)	28.41(0.7863)	28.52 (0.7883)	28.82(0.7976)	28.69 (0.7937)	**28.84** (**0.7995**)
	×4	25.96(0.6675)	26.82(0.7087)	26.75(0.7054)	26.90(0.7101)	27.06 (0.7167)	**27.29** (0.7251)	27.14 (0.7200)	27.28 (**0.7272**)

extraction in CompNet, can also prevent the gradient vanishing problem [32] for the first and the $(N/2)th$ layers, where N is the depth of the network. The backpropagation relations corresponding to the first and $(N/2)th$ layers, which are used in the reconstruction process, can be formulated as

$$\begin{aligned}\nabla_{W_{\frac{N}{2}}} L &= \frac{\partial L}{\partial W_N}\prod_{i=\frac{N}{2}+1}^{N}\frac{\partial W_i}{\partial W_{i-1}} + \frac{\partial L}{\partial W_N}\frac{\partial W_N}{\partial W_{\frac{N}{2}}} \\ &= \frac{\partial L}{\partial W_N}\Big(\prod_{i=\frac{N}{2}+1}^{N}\frac{\partial W_i}{\partial W_{i-1}} + \frac{\partial W_N}{\partial W_{\frac{N}{2}}}\Big) \\ \nabla_{W_1} L &= \frac{\partial L}{\partial W_N}\prod_{i=2}^{N}\frac{\partial W_i}{\partial W_{i-1}} + \frac{\partial L}{\partial W_N}\frac{\partial W_N}{\partial W_{\frac{N}{2}}}\prod_{i=2}^{\frac{N}{2}}\frac{\partial W_i}{\partial W_{i-1}} \\ &\quad + \frac{\partial L}{\partial W_N}\frac{\partial W_N}{\partial W_1} \\ &= \frac{\partial L}{\partial W_N}\Big(\prod_{i=2}^{N}\frac{\partial W_i}{\partial W_{i-1}} + \frac{\partial W_N}{\partial W_{\frac{N}{2}}}\prod_{i=2}^{\frac{N}{2}}\frac{\partial W_i}{\partial W_{i-1}} + \frac{\partial W_N}{\partial W_1}\Big)\end{aligned} \tag{6}$$

where L is the loss between the ground truth samples and the corresponding estimated high resolution samples and W_i represents the weight parameters of the ith layer. It should be noted that the value of the expression for each product in (6) is prone to becoming very small. However, the presence of the additive terms in (6) arising from the skip connection in CompNet should help to resolve the problem of vanishing gradient that is generally faced by other networks.

IV. EXPERIMENTAL RESULTS

In this section, the results of various experiments that are conducted using CompNet are presented and analyzed. To start with, the training and evaluation datasets are introduced. The performance of CompNet in terms of PSNR and SSIM [36] is presented, analyzed and compared with that of other schemes. In addition, the effects of the hyper parameters, width and depth, in CompNet are investigated. Moreover, the impact of dilated kernels [37] to increase the receptive field of the CompNet is explored. Finally, the effect of negative features on the network performance through the use of the exponential linear unit (ELU) [38] is examined.

A. DATASET

In our experiments, we use *BSD200* [27] and *91 images* [2] that have, respectively, 200 and 91 images, as a training dataset. After data augmentation, a total number of 151815 sub-images are generated from this dataset and utilized as the the training samples. For the purpose of evaluation, *Set5* [39], *Set14* [40] and *BSD100* [27] are employed with the upscaling factors of 2, 3 and 4. The training dataset is divided into batches each of size 64 (with the exception of the last batch), Thus each epoch has 2373 iterations (backpropagations). The number of epochs is set as 80 for the sake of consistency, when comparing the proposed with other schemes. Also, the weight decay parameter is set as 10^{-4}.

All of the deep learning tasks are implemented using Keras [41] that is backended by TensorFlow package [42]. The training procedure for CompNet is conducted on a machine with Intel Core i7 CPU @4.2 GHz, 16 GB installed memory and GPU Nvidia Titan X (Pascal).

B. PERFORMANCE OF COMPNET AND COMPARISON WITH THE STATE-OF-THE-ARTS

In this part, the objective evaluation in terms of PSNR and the subjective evaluation in terms of SSIM are carried out between the ground truth image and the estimated high resolution image. The results of CompNet and five state-of-the-art schemes namely, A+ [43], RFL [44], SRCNN, cascaded SCN (CSCN) and VDSR, are given in Table 1. It is seen from the results of this table that in almost all the cases, CompNet outperforms all of the state-of-the-art schemes. In some cases, improvement in the performance provided by CompNet is quiet significant. For instance, in the case of *Set14* test set (upscaled by 3), CompNet yields a PSNR value that is 0.29 dB higher than that given by VDSR along with the similarity measure that %0.54 higher.

In order to compare the visual quality of the images estimated by using the different methods, we provide in Fig. 2 the high resolution *Lena* image upscaled by a factor of 3.

It is seen from this figure that the high resolution images obtained by other methods, with the exception of VDSR and CompNet, suffer from ringing effects around some of the edges. It should be noted that although it is difficult to notice the difference between the visual quality of the images restored by using VDSR and CompNet, nevertheless CompNet provides a performance superior to that provided by VDSR in terms of both the objective and subjective measures.

FIGURE 2. Visual comparison between CompNet and the other state-of-the-art methods when applied to the *Lena* image with upscaling factor 3. (a) Ground truth. (b) A+ (PSNR = 33.55, SSIM = 0.8862). (c) SRCNN (PSNR = 33.70, SSIM = 0.8878). (d) CSCN (PSNR = 33.70, SSIM = 0.8879). (e) VDSR (PSNR = 33.99, SSIM = 0.8915). (f) CompNet (PSNR = 34.02, SSIM = 0.8920).

Fig. 3 shows the results of the super resolution carried out by CompNet on the *Butterfly* image interpolated by factors of 2, 3 and 4. It is seen from this figure that all the restored images have a very good visual quality. The super resolved image upscaled by a factor 2 is exceptionally closer to the ground truth. The reason for this is that the downsampling of an image with a lower scaling factor represents smaller loss of information and thus, the restored image should be more similar to the original one. However, in the case of higher upscaling factors, CompNet still provides visually very good results.

It is worth noting that more recently, some deep learning schemes that have attempted to improve the performance of single image super resolution by employing rather very large

FIGURE 3. *Butterfly* **images resulting from CompNet. (a) Ground truth. (b) Bicubic downsampling by a factor of 3. (c) Bicubic upsampling by a factor of 3. (d) Super resolved by a factor of 2. (e) Super resolved by a factor of 3. (f) Super resolved by a factor of 4.**

number of parameters in their networks have been proposed. For instance, the average performance of the newly proposed EDSR [45] and SRResNet [46] in terms of PSNR (SSIM) are, respectively, 32.46 dB (0.8968) and 32.05 dB (0.8910) for the Set5 images [39] with the upscaling factor 4. However, this superior performance of these networks is achieved at the expense of the numbers of parameters that are 16 and 2 times higher than that used by the proposed CompNet, whose complexity is discussed in the next subsection.

C. COMPLEXITY ANALYSIS OF COMPNET

With the default settings, CompNet is composed of 19 layers each of width 64 followed by a layer of width one. Now, we consider two separate variations in the default settings of the hyper parameters of CompNet. In the first one, we reduce the depth of CompNet from 20 to 16. The resulting network is referred to as reduced-depth CompNet (*RD CompNet*), whereas, in the second one, the width of CompNet is reduced from 64 to 32 and refer the resulting network as reduced-width CompNet (*RW CompNet*). Table 2 gives the total number of parameters for the three settings of CompNet as well as that for SRCNN and VDSR. It is seen from this table that the number of parameters for SRCNN is the lowest. However, it is not a deep network. On the other hand,

TABLE 2. Complexity of Various Super Resolution Schemes.

Method	Number of Parameters
SRCNN (Reproduced)	57281
VDSR (Reproduced)	665921
ComNet	673605
RD CompNet	524869
RW CompNet	170405

CompNet with the two new settings has a complexity lower than that of VDSR with RW CompNet having a considerably lower complexity. Fig. 4 gives the PSNR value as a function of the number of epochs. It is seen from this figure that CompNet provides a performance that is superior to that of RD CompNet or RW CompNet. However, the performance of the two latter networks are not substantially different, thus indicating the robustness of CompNet with respect to its width and depth. It is worth noting that the number of parameters of RW CompNet is substantially lower than that of CompNet with only a modest decrease in the performance.

We now run another experiment using the proposed network in which the number of parameters is reduced from 673605 to 636421 by removing one of the nonlinear mapping layers (i.e., layer 19). We have applied the network

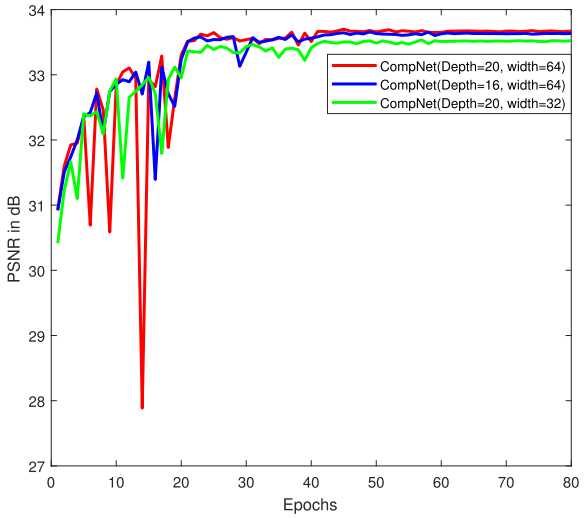


FIGURE 4. CompNet convergence for different settings of hyperparameters on *Set 5* with upscaling factor 3.

with reduced number of parameters on the three benchmark datasets. The results obtained from our network and that obtained by applying VDSR, which uses 665921 parameters, are shown in Table 3. It is seen from this table that the proposed network still yields a performance superior to that given by VDSR. It is clear from the results of this experiment that the performance gain of the proposed network over VDSR cannot be simply attributed to the use of a larger number of parameters, but rather to the use of an appropriate composition of the features in reconstructing the residue.

TABLE 3. PSNR (SSIM) results of CompNet with 19 layers and VDSR. The bolded values are the best in the comparison.

Dataset	scaling	VDSR	CompNet
Set5	×2	37.53 (0.9587)	**37.54** (**0.9594**)
	×3	**33.66**(0.9213)	33.64(**0.9215**)
	×4	**31.35**(**0.8838**)	31.29 (0.8829)
Set14	×2	33.03(0.9124)	**33.25** (**0.9149**)
	×3	29.77(0.8314)	**30.05** (**0.8367**)
	×4	28.01(0.7674)	**28.23** (**0.7730**)
BSD100	×2	31.90(0.8960)	31.90 (**0.8971**)
	×3	28.82(0.7976)	**28.83** (**0.7994**)
	×4	**27.29** (0.7251)	27.27 (**0.7271**)

D. EFFECT OF A DILATED KERNEL ON COMPNET

By increasing the effective receptive field of the network, a large range of information is included for restoring the high resolution image, and therefore, one could expect to improve the performance. One way of increasing the effective receptive field of a network is to increase the kernel size. However, this approach would increase the complexity of the network. Dilated convolution is a technique that would increase the receptive field of the network without increasing the network complexity. In a dilated convolution, the kernel is zeropadded to extract the long range information that can be utilized for the problem of image super resolution. The convolution using a dilated kernel, $h[m, n]$ with the dilation rate d, can be expressed as

$$y[m,n] = \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} x[m-dk, n-dl]h[k,l] \qquad (7)$$

where $x[m, n]$ is a two-dimensional signal to be convolved and $y[m, n]$ is the signal resulting from the dilated convolution.

TABLE 4. PSNR (SSIM) Values of CompNet Using Dilated Convolution with $d = 2$, when Applied to *Set5* Images.

Upscaling factor	Undilated	Dilated
2	**37.58** (**0.9596**)	37.41 (0.9588)
3	33.67 (**0.9219**)	**33.71** (**0.9219**)
4	**31.35** (0.8833)	31.31 (**0.8835**)

In our experiment on the impact of dilated convolution, we chose the dilation rate d to be 2 and the test datasets *Set5*. Table 4 gives the PSNR and SSIM results from the undilated and dilated convolutions with upscaling factors 2, 3 and 4. It is seen from this table that in the case of the upscaling factor of 2, inclusion of long-range information through the dilated convolution is detrimental to the network performance. However, for larger upscaling factors, dilation does not affect its performance.

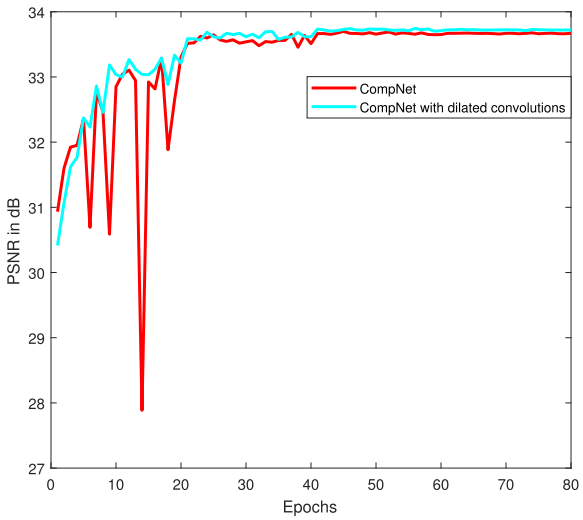
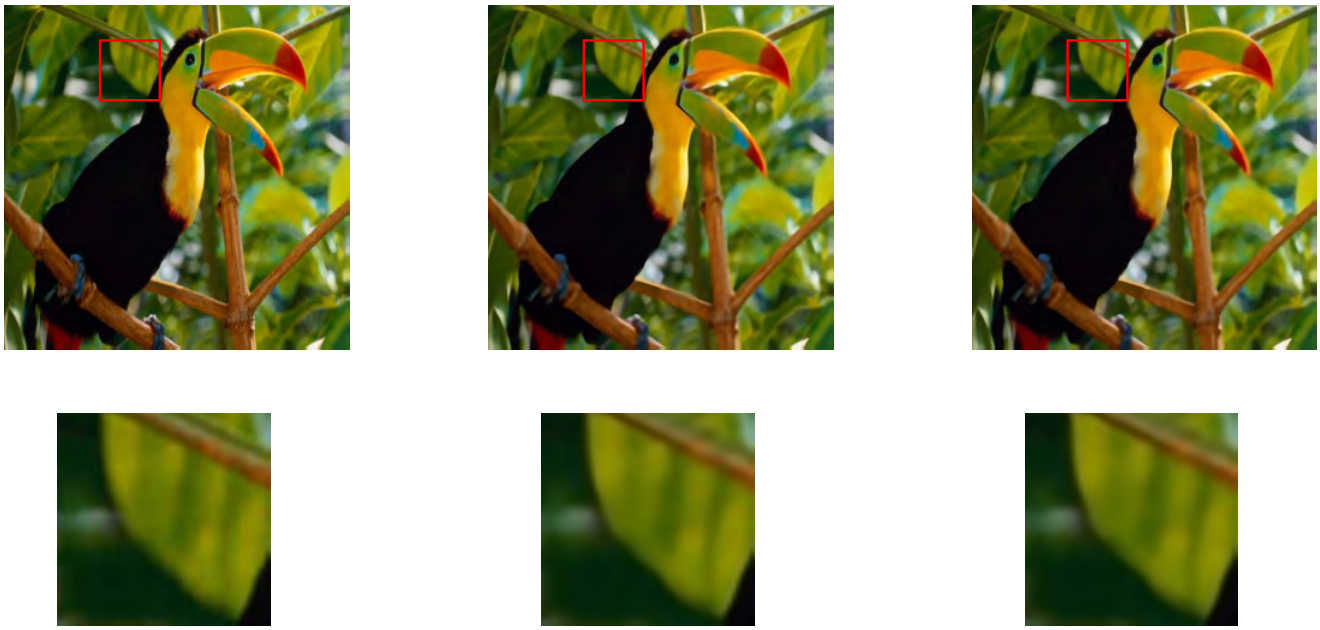


FIGURE 5. CompNet with $d = 2$ and without a dilated kernel on *Set 5* with upscaling 3.

Figure 5 shows the convergence performance of CompNet when applied on *Set5* images upscaled by a factor of 3 with and without dilated kernel. It is seen from this figure that the use of a dilated convolution does not provide any meaningful advantage over the undilated one in terms of either the convergence rate or the final PSNR value.

E. EFFECT OF USING NEGATIVE FEATURES

In our experiments on the proposed CompNet, we have so far used ReLU as the activation function. ReLU obviously does



(a) (b) (c)

FIGURE 6. *Bird* images resulting from CompNet using different activation functions. (a) Ground truth. (b) Using ELU activation function. (c) Using ReLU activation function.

not allow the negative features to be propagated through the network. We now examine the performance of the network in the presence of negative features as well as by replacing ReLU with the ELU function defied by

$$f(z) = \begin{cases} z & z \geq 0 \\ e^z - 1 & z < 0 \end{cases} \tag{8}$$

where z is the input to the activation unit. As can be seen from the above equation, ELU scales negative input values instead suppressing them to zero.

Comparative results in terms of PSNR and SSIM of CompNet with the ELU and ReLU activation functions on the *Set5* images and the illustration of the restored *Bird* image with the upscaling factor of 3 are depicted in Table 5 and Fig. 6, respectively.

TABLE 5. PSNR (SSIM) Values of CompNet Using ReLU and ELU Activation Functions, when Applied to *Set5* Images.

Upscaling factor	ReLU	ELU
2	**37.58 (0.9596)**	37.42 (0.9584)
3	**33.67 (0.9219)**	33.55 (0.9200)
4	**31.35 (0.8833)**	31.22 (0.8803)

It is seen from the PSNR and SSIM values given in the table that the negative features are not vital in improving the performance of CompNet. This conclusion is also supported by the visual quality of the *Bird* images restored using ReLU and ELU activation functions.

V. CONCLUSION

In this work, a residual deep neural network architecture, called CompNet, has been proposed for the problem of single image super resolution. In traditional neural network, the features produced by the layers become increasingly more sparse, as they become more deep. The objective of this work has been to compose a residual signal that is less sparse and a better representative of the different types features produced by the network layers. In the proposed CompNet, the residual signal is composed by taking into account the features produced by the initial, middle and last layer of the network. This particular composition of the residual signal contains a richer information about the estimated high resolution image without unduly increasing the network complexity. Extensive experiments have been carried out on CompNet using three benchmark datasets and the results in terms of PSNR, SSIM and the visual quality of the restored high resolution images have been obtained. It has been demonstrated that the proposed residual deep network outperforms the state-of-the-art schemes.

REFERENCES

[1] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, May 2003.
[2] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.
[3] K.-W. Hung and W.-C. Siu, "Single-image super-resolution using iterative Wiener filter based on nonlocal means," *Signal Process., Image Commun.*, vol. 39, pp. 26–45, Nov. 2015.
[4] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[5] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016, ch. 9.
[6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. CVPR*, 2016, pp. 770–778.
[7] B. Li and Y. He, "An improved ResNet based on the adjustable shortcut connections," *IEEE Access*, vol. 6, pp. 18967–18974, 2018.
[8] O. Russakovsky *et al.*, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.
[9] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. CVPR*, 2015, pp. 3431–3440.
[10] S.-J. Lee, T. Chen, L. Yu, and C.-H. Lai, "Image classification based on the boost convolutional neural network," *IEEE Access*, vol. 6, pp. 12755–12768, 2018.
[11] L. Zhang *et al.*, "Improving semantic image segmentation with a probabilistic superpixel-based dense conditional random field," *IEEE Access*, vol. 6, pp. 15297–15310, 2018.
[12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. NIPS*, 2012, pp. 1097–1105.
[13] M. D. Zeiler *et al.*, "On rectified linear units for speech processing," in *Proc. ICASSP*, 2013, pp. 3517–3521.
[14] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2015.
[15] C. Dong, C. C. Loy, and X. Tang, "Accelerating the super-resolution convolutional neural network," in *Proc. ECCV*, 2016, pp. 391–407.
[16] D. Liu, Z. Wang, B. Wen, J. Yang, W. Han, and T. S. Huang, "Robust single image super-resolution via deep networks with sparse prior," *IEEE Trans. Image Process.*, vol. 25, no. 7, pp. 3194–3207, Jul. 2016.
[17] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proc. ICML*, 2010, pp. 399–406.
[18] D. Liu, Z. Wang, N. Nasrabadi, and T. Huang, "Learning a mixture of deep networks for single image super-resolution," in *Proc. ACCV*, 2016, pp. 145–156.
[19] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proc. CVPR*, 2016, pp. 1646–1654.
[20] K. Simonyan and A. Zisserman. (2014). "Very deep convolutional networks for large-scale image recognition." [Online]. Available: https://arxiv.org/abs/1409.1556
[21] Y. Liang, Z. Yang, K. Zhang, Y. He, J. Wang, and N. Zheng. (2017). "Single image super-resolution via a lightweight residual convolutional neural network." [Online]. Available: https://arxiv.org/abs/1703.08173
[22] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
[23] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.
[24] A. Esmaeilzehi and H. A. Moghaddam, "Nonparametric kernel sparse representation-based classifier," *Pattern Recognit. Lett.*, vol. 89, no. 4, pp. 46–52, 2017.
[25] T. Peleg and M. Elad, "A statistical prediction model based on sparse representations for single image super-resolution," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2569–2582, Jun. 2014.
[26] W. Yang *et al.*, "Deep edge guided recurrent residual learning for image super-resolution," *IEEE Trans. Image Process.*, vol. 26, no. 12, pp. 5895–5907, Dec. 2017.
[27] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. ICCV*, 2001, pp. 416–423.
[28] Y. Wang, L. Wang, H. Wang, and P. Li. (2016). "End-to-end image super-resolution via deep and shallow convolutional networks." [Online]. Available: https://arxiv.org/abs/1607.07680
[29] J. W. Woods, *Multidimensional Signal, Image, and Video Processing and Coding*, 2nd ed. New York, NY, USA: Academic, 2001.
[30] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger. (2018). "Densely connected convolutional networks." [Online]. Available: https://arxiv.org/abs/1608.06993
[31] S. Ioffe and C. Szegedy. (2015). "Batch normalization: Accelerating deep network training by reducing internal covariate shift." [Online]. Available: https://arxiv.org/abs/1502.03167
[32] R. Pascanu, T. Mikolov, and Y. Bengio. (2013). "On the difficulty of training recurrent neural networks." [Online]. Available: https://arxiv.org/abs/1211.5063
[33] C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals. (2017). "Understanding deep learning requires rethinking generalization." [Online]. Available: https://arxiv.org/abs/1611.03530
[34] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. ICCV*, 2015, pp. 1026–1034.
[35] T. Doza. *Incorporating Nesterov Momentum Into Adam*. Accessed: Oct. 9, 2018. [Online]. Available: https://web.stanford.edu/~tdozat/files/TDozat-CS229-Paper.pdf
[36] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[37] F. Yu and V. Koltun. (2016). "Multi-scale context aggregation by dilated convolutions." [Online]. Available: https://arxiv.org/abs/1511.07122
[38] D.-A. Clevert, T. Unterthiner, and S. Hochreiter. (2016). "Fast and accurate deep network learning by exponential linear units (ELUs)." [Online]. Available: https://arxiv.org/abs/1511.07289
[39] M. Bevilacqua, A. Roumy, C. Guillemot, and M.-L. Alberi-Morel, "Low-complexity single-image super-resolution based on nonnegative neighbor embedding," in *Proc. BMVC*, 2012, pp. 135.1–135.10.
[40] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," in *Curves and Surfaces*. Heidelberg, Germany: Springer, 2012.
[41] F. Chollet. (2015). *Keras*. Accessed: Oct. 9, 2018. [Online]. Available: https://github.com/keras-team/keras
[42] M. Abadi *et al.* (2015). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Accessed: Oct. 9, 2018. [Online]. Available: https://www.tensorflow.org
[43] R. Timofte, V. De Smet, and L. Van Gool, "A+: Adjusted anchored neighborhood regression for fast super-resolution," in *Proc. ACCV*, 2014, pp. 111–126.
[44] S. Schulter, C. Leistner, and H. Bischof, "Fast and accurate image upscaling with super-resolution forests," in *Proc. CVPR*, 2015, pp. 3791–3799.
[45] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," in *Proc. CVPR*, 2017, pp. 1132–1140.
[46] C. Ledig *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. CVPR*, 2017, pp. 4681–4690.

ALIREZA ESMAEILZEHI received the B.Sc. degree in biomedical (bioelectrical) engineering from the University of Isfahan, Isfahan, Iran, in 2013, and the M.Sc. degree in biomedical (bioelectrical) engineering from the K. N. Toosi University of Technology, Tehran, Iran, in 2015. He is currently pursuing the Ph.D. degree in electrical and computer engineering with Concordia University, Montreal, QC, Canada, under the supervision of Dr. M. O. Ahmad and Dr. M.N.S. Swamy. His research interests include the areas of machine and deep learning, image processing, computer vision, and compressive sensing.

M. OMAIR AHMAD (S'69–M'78–SM'83–F'01) received the B.Eng. degree in electrical engineering from Sir George Williams University, Montreal, QC, Canada, and the Ph.D. degree in electrical engineering from Concordia University, Montreal. From 1978 to 1979, he was a Faculty Member with New York University College, Buffalo, NY, USA. In 1979, he joined the faculty of Concordia University as an Assistant Professor of computer science. Subsequently, he joined the Department of Electrical and Computer Engineering, Concordia University, where he was the Chair with the Department from 2002 to 2005, and is currently a Professor. He was a Founding Researcher of Micronet, a Canadian Network of Centers of Excellence, from 1990 to 2004. He is also the Concordia University Research Chair (Tier I) in multimedia signal processing. He has authored in the area of signal processing and holds four patents. His current research interests include the areas of image and speech processing, biomedical signal processing, watermarking, biometrics, video signal processing and object detection and tracking, deep learning techniques in signal processing, and fast signal transforms and algorithms. In 1988, he was a member of the Admission and Advancement Committee of the IEEE. He was a recipient of numerous honors and awards, including the Wighton Fellowship from the Sandford Fleming Foundation, an induction to Provosts Circle of Distinction for Career Achievements, and the Award of Excellence in Doctoral Supervision from the Faculty of Engineering and Computer Science, Concordia University. He was a Guest Professor with Southeast University, Nainjing, China, and the Local Arrangements Chairman of the 1984 IEEE International Symposium on Circuits and Systems. He has served as the Program Co-Chair for the 1995 IEEE International Conference on Neural Networks and Signal Processing, the 2003 IEEE International Conference on Neural Networks and Signal Processing, and the 2004 IEEE International Midwest Symposium on Circuits and Systems. He was a General Co-Chair of the 2008 IEEE International Conference on Neural Networks and Signal Processing. He is the Chair of the Montreal Chapter IEEE Circuits and Systems Society. He was an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS PART I: FUNDAMENTAL THEORY AND APPLICATIONS from 1999 to 2001.

M.N.S. SWAMY (S'59–M'62–SM'74–F'80) received the B.Sc. degree (Hons.) in mathematics from the University of Mysore, Mysore, India, in 1954, the Diploma degree in electrical communication engineering from the Indian Institute of Science, Bangalore, India, in 1957, and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Saskatchewan, Saskatoon, SK, Canada, in 1960 and 1963, respectively. He was conferred with the title of Honorary Professor by the National Chiao Tong University, Hsinchu, Taiwan, in 2009. He is currently a Research Professor with the Department of Electrical and Computer Engineering, Concordia University, Montreal, QC, Canada, where he served as the Founding Chair of the Department of Electrical Engineering from 1970 to 1977, and the Dean of engineering and computer science from 1977 to 1993. During that time, he developed the faculty into a research-oriented one from what was primarily an undergraduate faculty. Since 2001, he has been the Concordia Chair (Tier I) in signal processing. He has also taught with the Department of Electrical Engineering, Technical University of Nova Scotia, Halifax, NS, Canada, the University of Calgary, Calgary, AB, Canada, and the Department of Mathematics, University of Saskatchewan. He has published in the areas of number theory, circuits, systems, and signal processing, and holds five patents. He has co-authored nine books and five book chapters. He was a Founding Member of Micronet, Ottawa, Canada, a Canadian Network of Centers of Excellence from 1990 to 2004, and also its Coordinator of Concordia University. He is a fellow of the Institute of Electrical Engineers, U.K, the Engineering Institute of Canada, the Institution of Engineers, India, and the Institution of Electronic and Telecommunication Engineers, India. He was inducted in 2009 to the Provosts Circle of Distinction for career achievements. He is a recipient of many IEEE-CAS Society awards, including the Education Award in 2000, the Golden Jubilee Medal in 2000, and the 1986 Guillemin-Cauer Best Paper Award. He served as a Program Chair for the 1973 IEEE Circuits and Systems (CAS) Symposium, a General Chair of the 1984 IEEE CAS Symposium, a Vice Chair of the 1999 IEEE CAS Symposium, and a member of the Board of Governors of the CAS Society. He served as the Editor- in-Chief for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I from 1999 to 2001, and an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS from 1985 to 1987. He has served the IEEE in various capacities such as the President Elect in 2003, President in 2004, Past-President in 2005, and Vice President (publications) from 2001 to 2002, Vice President in 1976. He has been the Editor-in-Chief of the journal *Circuits, Systems and Signal Processing* (CSSP) since 1999. Recently, CSSP has instituted the Best Paper Award in his name.