Single Image Super Resolution Algorithms: A Survey and Evaluation

Ruaa Adeeb Abdulmunem Al-falluji*, Aliaa Abdel-Halim Youssif, Shawkat K. Guirguis*****

** lecturer ,Babylon University, Babylon, Iraq.*
*** Dean, faculty of Computers and Information, Helwan University ,Cairo,Egypt.*
**** Vice Dean for Graduate Studies and Research ,Institute of Graduate Studies & Research (IGSR), Alexandria University, Alexandria,Egypt.*

***Abstract*— Image processing sub branch that specifically deals with the improvement, of images and videos, resolution without compromising the detail and visual effect but rather enhances the two, is known as Super Resolution. Multiple (multiple input images and one output image) or single (one input and one output) low resolution images are converted to high resolution. Single image super resolution algorithms are more practical since multiple images are not always available. The paper presents a survey of recent single image super resolution methods that are based on the use of external database to predict the values of missing pixels in high resolution image.**

Keywords—Super resolution, Sparese dictionary, Random forest, Convolution neural network

I. Introduction

Nowadays a lot of research is being done on image processing techniques that can improve the visual quality of images and videos. The need of such techniques stem from progress in technology as well as people desire for high quality multimedia content. The past few decades has seen phenomenal advancement in the viewing equipment such as PDPs, LCDs and LEDs(Light emitting displays) which display a picture with high spatial and temporal parameters making it crystal clear and visually pleasing.

But despite the demand and interest in high-resolution content they are not always available due to several reasons that include down sampling for the sake of bandwidth limitations, different type of noise, different compression techniques and so on. Another major reason is the quality of the imaging equipment generally available

There are many other reasons, which are not entirely related to user experience, that are driving force behind development of such algorithms. Some of these include video surveillance [1], remote sensing [2], medical and military applications. In fact any application that requires zoom in operation or recognition needs to have super resolution as part of it to make results more understandable. Another major is the incompatibility between different video standards. To make them compatible with each other resolution conversion is a basic step.

Super resolution refers to a set of algorithms whose purpose is to increase the resolution of images and video. Upscaling and upsizing means the same thing but super resolution methods try to do so without sacrificing the detail and visual appearance of the images or video. Super resolutions main goal is to find the value of the missing pixels in the high resolution image.

Super resolution methods can further be classified into two classes: Multiple image super resolution methods [3, 4, 5, 6] and single image super resolution methods. Multiple image super resolution methods operate on the assumption that the multiple images that are being used for construction of the high resolution image are geometric transforms and misaligned versions of each other and by combining them in suitable manner will yield an image with more detail then contained in separate images. But it is not a very practical as multiple images are not always available. So single image super resolution methods are now worked on and are the subject of this survey. Over the years several single image super resolution algorithms have been developed. The survey is about example based super resolution algorithms. They can be divided into three main categories.

Internal database based Super resolution algorithms: A number of proposed algorithms [7, 8] make use of the natural image self-similarity property to construct scale space pyramids of low resolution, to match low resolution and high resolution pairs. Though the training patches contained in the local database are more relevant than that of external database, the number of local training patches is very limited. To overcome this several methods were used e.g. Singh et al[9] instead of using direct patches first transformed them into directional sub bands and found matches for them in sub pyramids. Huang et al.[10] on the other hand expanded the search space such that it now included both perspective and affine transform of patches. The major drawback of these methods is time needed for their computation, making them not suitable for real time problems.

External database based super resolution methods: the second type of example based super resolution methods use external images to try and find mapping between the high resolution and low resolution images. The algorithms use different supervised machine learning techniques such as nearest neighbor [11], manifold embedding [12, 13], kernel ridge regression [14], and sparse representation [15, 16, 17]. These methods also divide the database into clusters using kmeans [18], random forests [19] and sparse dictionary [20] and then try and find linear repressors for each cluster separately.

CNN based super resolution methods: the most recent techniques used by researchers to construct high resolution image using CNN's. CNN's optimize all the steps in super resolution process and try a nonlinear mapping function between low resolution and high resolution in image space. Several methods using CNN are proposed, the most recent being VSDR [21], DRSN [22] AND LapSRN [23].

II. ALGORITHMS

A. A+: Adjusted Anchored Neighborhood Regression [20]

To create the dictionary authors use the method proposed by Zeyde et al.[24] and ANR[25]. In the proposed method first optimization is performed over low resolution image patches in order to obtain a sparse dictionary, represented by Dl. The Dl is then used to reconstruct its corresponding high resolution image dictionary, represented by, Dh. This is done such that the coefficients from the low resolution image patch decompositions over Dl are the same as the corresponding coefficients in the high resolution image patch decompositions over Dh.

The commonly used methods for dictionary construction i.e. ANR and the other sparse coding approaches do not use the training samples for any purpose other than the dictionary construction, but in the proposed method they play a vital role as the neighborhood used for calculation of regression is also taken from the training images which is formulated in the training phase and is then used at test time. Regression used at training phase is the same as proposed in the ANR i.e ridge regression, the only difference between the two is that rather than using the atoms in the sparse dictionary as neighborhood, the neighborhood is defined in terms of the dense training i.e. rather than using the neighborhood of atoms, for which a unique version exists for each of the dictionary atom, a matrix containing samples from the training set that lie closest to the dictionary atom to which the input patch y is matched. The Nearest neighbor search is done based on euclidean distance which is calculated between the samples used for training and the anchor atom.

In order to have a vigorous repressor for each atom, the neighborhood of samples needs to be brought to unit l2 norm i.e. should be brought on the unit hyper sphere surface, or should be centered on the atom. To avoid any noise that can be generated due to the flatness of a patch, l2 normalization is not performed on the patches whose features values are below are certain threshold. In such cases the set of neighboring samples are used for calculating the local manifold on the unit hyper sphere, even if they are not all lying on the hyper sphere. But due to the algorithm used for low resolution feature extraction uniform scaling factor is always obtained between the low resolution feature and its High resolution image patch so when the low resolution image patch features are transferred to the hyper sphere by using the l2 normalization, the corresponding high resolution patches are also linearly scaled by multiplying them with the same factor so that the relation between low resolution and high resolution spaces remain the same. It was observed that the local manifold approximation calculated using regression will be better if the samples are closer to the anchoring atom. Thus, training samples equal to the neighborhood size are retrieved for each atom. As a result different atom centered neighborhoods share same samples. By doing so it is guaranteed that even in the extreme conditions such as where the number of training samples are small compared to the number of atoms or the number of samples related to each atom is small, the repressors are still learned robustly.

B. Transformed Self-Exemplars(SelfExSR)[10]

For a given low resolution image IL, the first step is to blur and subsample the image in order to get a down sampled image denoted by ID. The image I is then divided into patches. Each patch is then wrapped to a patch in the down sampled images, which is best match to the said patch, by the use of homographic transformation matrix. A modified version of PatchMatch [26] algorithm is used to calculate the parameters of the transformation matrix. The patchMatch algorithm use three cost measures namely scale, plane and appearance costs to estimate parameters using following steps.

Initialization: where the nearest neighbor filed is initialized using desired scale and zero displacement rather than random initialization.

Propagation: In this steps the parameters on the transformation are propagated for the matches of neighbors rather than matrix. By doing so source patch placement becomes invariant to affine transform.

Randomization: In this step a random search is done to check if a better solution exists.

After the transformation matrix is obtained a high resolution version of the source patch is calculated. Inverse transformation matrix is then applied on the high resolution patch to unwrap the patch in order to obtain self-exemplars which is the final high resolution version of the patch. The high resolution patch is then places in the space of low resolution patch to form the High resolution image. The whole process is repeated for all the patches of the low resolution image. After all the patches have been processed an iterative backpropogation algorithm [27] is applied in order to check that all the constraints of the low resolution image are satisfied in the high resolution image.

C. Super-Resolution using Random Forests(RF)[19]

Super resolution algorithms basically work by finding a mapping function between the low resolution image and high resolution image. Up until this point all proposed algorithms do so by using a slightly different versions of the couple's dictionary learning. But all these methods have the drawbacks of being slow and require sparse encoding. The algorithm proposed by Criminisi [28] suggest the use of random forest for learning the mapping function. Use of random forests enable use of parallization thus reducing time needed for processing. In the proposed algorithm data dependencies were generated using random forests. Random forests are groups of binary trees. All tress of random forests are trained independently of each other using 'N' number of training samples. A single tree works by splitting the training data by using a splitting function into uneven sets. The starting node is called the root node and end leaf node. Splitting starts at the root node and continues until a leaf node is reached i.e. no more splitting is possible or the predefined tree depth is reached. By using the leaf nodes created by different trees a subset of overlapping cells. This separation is then used to find the data dependences and using these each leaf node can be represent by a linear model. The final data dependent mapping matrix is computed by using the average of trees generated.

D. Super-Resolution Using Very Deep Convolutional Networks(VDSR)[21]

For high resolution image construction Simonyan and Zisserman [29] inspired very deep convolution network was used. 'D' number of layers were used in the CNN. First layer take the LR image input whereas the last layer reconstructs the high resolution image and contains a single 3x3x64 filter. Each layer except the first and last one has the same configuration. Each layer consists of 3x3x64 size 64 filters where 3x3 refers to the spatial region of the filter whereas 64 represents the feature maps. The input image of the network is the interpolated version of the low resolution image which is the same size as that of the output image. It was observed that convolutional neural network based methods work more accurately if they have domain knowledge available for modelling high resolution image.In the proposed algorithm it was observed that using residual images have several advantages such as fast convergence and superior performance i.e. they give better PSNR. One main problem that is faced by deep networks while in prediction stage is that feature maps get reduced with every iteration of convolution because super resolution make use of surrounding pixels to correctly predict the pixels. It is a better approach in a way that surrounding pixels always give better information about center pixel. But same is not possible for the boundary pixels so zeros were padded to the boundary to overcome this problem and gives better results. In training following approaches are used to get better results.

Residual Learning – Instead of trying to predict the whole input image residual image is predicted and represented by a loss calculation layer. The training continues until a certain Loss threshold is achieved

High learning rates for very deep networks – Setting high learning rates can cause exploding / vanishing gradient problems. So to overcome these adjustable gradient clippings are used. Gradient clipping refers to restricting individual gradients in the range between -θ and θ.

Multiscale Model – lastly multiscale model makes sharing of parameters possible between different scale factors instead of defining them separately for each factor.

To get final high resolution image all image details are predicted and then summed to the input image.

E. Deeply-Recursive Convolutional Network for Image Super-Resolution(DRCN)[22]

The proposed algorithm can be divided into three main parts namely embedding net, inference net and the reconstruction net. A low resolution image is passed to the embedding net as input. It can handle both grayscale and RGB images. The embedding layer converts the image to a set of feature maps. The features are represented in a way that are acceptable to the interference net hidden layers. Interference net is the main layer of the system that actually performs super resolution. Image is divided into regions and each region is analyzed by recursive single layer. At each iteration the size of the filter widens. After the final recursion the resultant feature maps represent the high resolution image. They need to be transformed back to the original image space. This step is done in the reconstruction net. Each of the sub nets used have single hidden layer and of these only one of them is recursive i.e the one in the inference net layer. All other are same as mutilayer perceptrons. As with any recursive model, the above model has its pros and cons. While recursion makes the system powerful and easy to implement the training itself is very difficult due to vanishing and exploding gradient effect. Exploding gradient means the exponent increase of gradient in the norm during the training which occurs due to the chained gradients multiplicative behavior. While vanishing gradients means the opposite of exploding gradients i.e. components that remain long term then reach zero really fast which cause the relation calculation between distant neighbors difficult. To overcome these issues two more layers were added to the model i.e Recursive-Supervision and skip connection layer. Recursive-Supervision layer monitors each recursion, as it was assumed that the same convolution occurs every time it can be assumed that same net can be used for reconstruction each time as well. Thus now all predictions are used for calculating output. Secondly a skip layer was added to the system between input layer and the reconstruction net. So now during recursion the input image goes directly from input layer to the reconstruction layer thus saving network capacity.

F. Deep Laplacian Pyramid Networks based Super-Resolution(LapSRN)[23]

The algorithm proposed is based on Laplacian pyramid. The system takes a low resolution image input rather than an interpolated version of the image, as done by other CNN based algorithms and systematically predicts, at each level of the pyramid, the residual image. The total number of levels of pyramid are equal to log2 S where S represents the scaling factor i.e for a scaling factor of 8 there will be 3 sub netwroks. The algorithm can be divided into two main parts.

Feature Extraction: Feature extraction is done at each level of the pyramid. 'D' convolutional layers followed by single transformed convolutional layer are used for the purpose of feature extraction. The feature extraction layer upsamples the feature map by a scaling factor of 2. The output generated by transposed convolutional layer is then feed to two different components, to a convolutional layer at the next pyramid level s+1 for further feature extraction and to a separate convolutional layer for reconstruction of residual image using feature generated at level s. To reduce the computational complexity of the system feature extraction is done a coarse resolution and then extract feature maps at finer resolution by using transposed convolutional layer. The feature maps at a level are shared with level above it which causes an increase in the non-linearity of the network.

In image reconstruction stage the input low resolution image is upscaled by a factor of 2 by the use of transposed convolutional layer. The layer is initialized using a bilinear kernel and use the same to optimize it with all other layer. Then element wise summation is used on upsclaed image and predicted residual image to generate a high resolution image. The high resolution image generated at each level of pyramid is then used as an input to the next layer of the reconstruction unit. Cascades of CNN's having same structure are used in the network at each level. In the algorithm proposed filters of size 3x3x64 are used by each convolutional layer. He at al proposed algorithm is used to initialize the convolutional filters. The transposed convolutional network used have a size of 4x4 and bilinear filter is used for initialization. Leaky rectified linear units having negative slope of 0.2 were used after each convolutional layer at the feature extraction stage. To keep feature maps size same as that of input image zeros are added at the boundaries.

III. Comparison

To compare the performance of above mentioned algorithms three measures, Peak signal to noise ratio (PSNR), Structural Similarity Index (SSI) and Information fidelity criteria (IFC), were used.

PSNR is the ratio between signal maximum power and the power of fidelity effecting noise of the signal Due to signal dynamic range the psnr is measured in decibel scale.

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}|I(i,j) - K(i,j)|^2$$

$$PSNR = 20.log_{10}(MAX)_I - 10.log_{10}(MSE)$$

Structural Similarity Index is a quantitative measure used to quantify the structural similarities between two images.

Information fidelity criteria is used to access the quality of an image based on Natural scene statistics. Mutual information between the test and reference image is quantified by the use of distortion model and the source signal, which signifies the perceptual quality of the image.

To test the algorithms 5 standard datasets i.e. SET14 [30] (number of images=14), SET5 [31] (number of images=5), MANGA109 [32] (number of images=109), BSDS100 [33] (number of images=100), and URBAN100 [10] (number of images=100) were used. Of the used datasets, SET5, SET14 and BSDS100 are the ones that are related to natural scenes while URBAN100 contains variety of different urban scenes images having a range of different frequency bands. Finally MANGA109 is a dataset of Japanese manga. The table below show the PSNR, SSI and IFC values for each algorithm for different output resolution.

Table 1. Quantitative Evaluation

	Scale	Set5	Set14	Bsds1000	Urban100	MANGA
		PSNR/SSI/IFC	PSNR/SSI/IFC	PSNR/SSI/IFC	PSNR/SSI/IFC	PSNR/SSI/IFC
A+	2	36.54 / 0.954 / 8.715	32.40 / 0.906 / 8.201	31.22 / 0.887 / 7.464	29.23 / 0.894 / 8.440	35.33 / 0.967 / 8.906
	4	30.30 / 0.859 / 3.260	27.43 / 0.752 / 2.961	26.82 / 0.710 / 2.564	24.34 / 0.720 / 3.218	27.02 / 0.850 / 3.177
	8	25.52 / 0.692 / 1.077	23.98 / 0.597 / 0.983	24.20 / 0.568 / 0.797	21.37 / 0.545 / 1.092	22.39 / 0.680 / 1.056
SelfExSR	2	36.49 / 0.954 / 8.391	32.44 / 0.906 / 8.014	31.18 / 0.886 / 7.239	29.54 / 0.897 / 8.414	35.78 / 0.968 / 8.721
	4	30.33 / 0.861 / 3.249	27.54 / 0.756 / 2.952	26.84 / 0.712 / 2.512	24.82 / 0.740 / 3.381	27.82 / 0.865 / 3.358
	8	25.52 / 0.704 / 1.131	24.02 / 0.603 / 1.001	24.18 / 0.568 / 0.774	21.81 / 0.576 / 1.283	22.99 / 0.718 / 1.244
RF	2	36.55 / 0.954 / 8.006	32.36 / 0.905 / 7.684	31.16 / 0.885 / 6.930	29.13 / 0.891 / 7.840	35.08 / 0.966 / 8.921
	4	30.15 / 0.853 / 3.135	27.33 / 0.748 / 2.853	26.75 / 0.707 / 2.455	24.20 / 0.711 / 3.000	26.80 / 0.840 / 3.055
	8	25.36 / 0.677 / 0.985	23.88 / 0.588 / 0.910	24.13 / 0.562 / 0.741	21.27 / 0.535 / 0.978	22.27 / 0.668 / 0.968
VDSR	2	37.53 / 0.958 / 8.190	32.97 / 0.913 / 7.878	31.90 / 0.896 / 7.169	30.77 / 0.914 / 8.270	37.16 / 0.974 / 9.120
	4	31.35 / 0.882 / 3.496	28.03 / 0.770 / 3.071	27.29 / 0.726 / 2.627	25.18 / 0.753 / 3.405	28.82 / 0.886 / 3.664
	8	25.72 / 0.711 / 1.123	24.21 / 0.609 / 1.016	24.37 / 0.576 / 0.816	21.54 / 0.560 / 1.119	22.83 / 0.707 / 1.138
DRCN	2	37.63 / 0.959 / 8.326	32.98 / 0.913 / 8.025	31.85 / 0.894 / 7.220	30.76 / 0.913 / 8.527	37.57 / 0.973 / 9.541
	4	31.53 / 0.884 / 3.502	28.04 / 0.770 / 3.066	27.24 / 0.724 / 2.587	25.14 / 0.752 / 3.412	28.97 / 0.886 / 3.674
	8	-	-	-	-	-
LapSRN	2	37.52 / 0.959 / 9.010	33.08 / 0.913 / 8.505	31.80 / 0.895 / 7.715	30.41 / 0.910 / 8.907	37.27 / 0.974 / 9.481
	4	31.54 / 0.885 / 3.559	28.19 / 0.772 / 3.147	27.32 / 0.728 / 2.677	25.21 / 0.756 / 3.530	29.09 / 0.890 / 3.729
	8	26.14 / 0.738 / 1.302	24.44 / 0.623 / 1.134	24.54 / 0.586 / 0.893	21.81 / 0.581 / 1.288	23.39 / 0.735 / 1.352

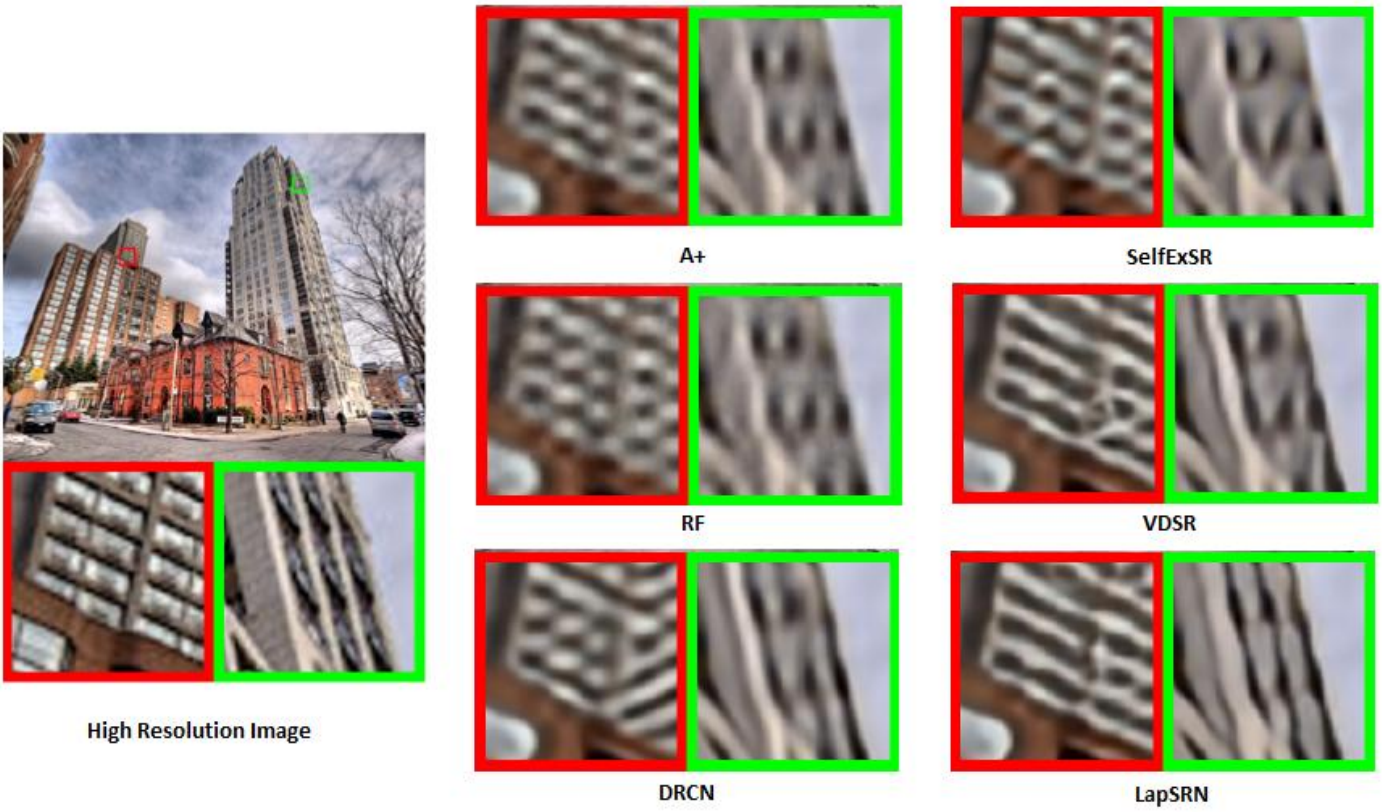


Figure 1. Visual Comparison: High Resolution Image: Ground Truth, A+ Result, SelfExSR Result, RF Result, VDSR Result, DRCN Result, LapSRN Result

Figure 1 shows the visual comparison of all the algorithms, on a scale of 4X, using an image from URBAN100 dataset.

Table 2 shows the frames per second comparison of the mentioned algorithm using a 3.4 GHz Intel i7 CPU (64G RAM) and Nvidia Titan X GPU (12G Memory).

Table 2. Frames per Second

	Scale	Set5	Set14	Bsds1000	Urban100	MANGA
A+	2	1.12	0.52	0.74	0.15	0.1
	4	2.86	1.62	2.43	0.49	0.41
	8	5.79	2.84	4.31	0.80	0.64
SelfExSR	2	0.02	0.01	0.01	0	0
	4	0.04	0.02	0.03	0	0
	8	0.03	0.01	0.02	0	0
RFL	2	0.65	0.45	0.52	0.13	0.15
	4	1.97	1.21	1.64	0.42	0.34
	8	2.54	1.61	2.25	0.47	0.33
VDSR	2	11.01	6.46	10	2.12	1.71
	4	10.71	6.59	9.91	2.15	1.76
	8	10.58	6.50	10.31	2.15	1.77
DRCN	2	0.70	0.37	0.59	0.10	0.08
	4	0.80	0.37	0.59	0.10	0.08
	8	-	-	-	-	-
LapSRN	2	30.20	40	97.36	16.81	85.32
	4	25.49	25.46	54.35	12.40	47.63
	8	24.09	23.40	50.44	10.54	33.09

From the tables and figure above it is apparent that not only LapSRN provide the best reconstruction but also can be used in real time application making it the best algorithm among the discussed algorithms

IV. Conclusion

The paper contains a survey of recent super resolution algorithms that are primarily based on example learning. Example learning methods are a lot accurate than self-patch reliant methods as the number of samples available for matching are more. Among the presented algorithms it was noticed that convolutional network based methods were a lot accurate than any other methods. The only drawback of the CNN's methods was the speed of the system that made them unsuitable for real time applications. The problem was solved by LapSRN algorithm which was found to be the most efficient both in terms of performance and efficiency, when applied on set of standard benchmark datasets.

References

[1] A. Gehani and J. H. Reif, "Super-Resolution Video Analysis for Forensic Investigations,"IFIP WG 11.9 International Conference on Digital Forensics, ser. IFIP, vol. 242. Springer, 2007, pp. 281_299.

[2] A. J. Tatem, H. G. Lewis, P. M. Atkinson, and M. S. Nixon, "Super-resolution target identification from remotely sensed images using a Hop_eld neural network," IEEE Transactionson Geoscience and Remote Sensing, vol. 39, no. 4, pp. 781_796, Apr. 2001.

[3] S. Borman and R. Stevenson, "Super-Resolution from Image Sequences - A Review", in Midwest Symposium on Circuits and Systems, Notre Dame, IN, USA, 8 1998, pp. 374_378.

[4] S. C. Park, M. K. Park, and M. G. Kang,"Super-Resolution Image Reconstruction: A Technical Overview", IEEE Signal Processing Magazine, vol. 20, no. 3, pp. 21_36, 5 2003.

[5] C. Mancas-Thillou and M. Mirmehdi, "An Introduction to Super-Resolution Text,"in Digital Document Processing, ser. Advances in Pattern Recognition. Springer London,2007, pp. 305_327.

[6] J. Tian and K.-K. Ma, "A survey on super-resolution imaging", Signal, Image and Video Processing (SIViP), vol. 5, no. 3, pp. 329_342, 2011.

[7] G. Freedman and R. Fattal. "Image and video upscaling from local self-examples". ACM TOG (Proc. of SIGGRAPH)30(2):12, 2011.

[8] D. Glasner, S. Bagon, and M. Irani. "Super-resolution from a single image". In ICCV, 2009.

[9] A. Singh and N. Ahuja. "Super-resolution using sub-band self-similarity". In ACCV, 2014.

[10] J.-B. Huang, A. Singh, and N. Ahuja. "Single image superresolution from transformed self-exemplars". In CVPR, 2015.

[11] W. T. Freeman, T. R. Jones, and E. C. Pasztor. "Examplebased super-resolution". IEEE, Computer Graphics and Applications, 22(2):56–65, 2002

[12] M. Bevilacqua, A. Roumy, C. Guillemot, and M. L. "Alberi- Morel. Low-complexity single-image super-resolution based on nonnegative neighbor embedding". In BMVC, 2012.

[13] H. Chang, D.-Y. Yeung, and Y. Xiong. "Super-resolution through neighbor embedding." In CVPR, 2004.

[14] K. I. Kim and Y. Kwon. "Single-image super-resolution using sparse regression and natural image prior." TPAMI, 32(6):1127–1133, 2010.

[15] J. Yang, J. Wright, T. Huang, and Y. Ma. "Image superresolution as sparse representation of raw image patches". In CVPR, 2008.

[16] J. Yang, J. Wright, T. S. Huang, and Y. Ma. "Image superresolution via sparse representation". TIP, 19(11):2861–2873, 2010.

[17] R. Zeyde, M. Elad, and M. Protter. "On single image scale-up using sparse-representations". In Curves and Surfaces. 2010.

[18] C.-Y. Yang and M.-H. Yang. "Fast direct super-resolution by simple functions". In ICCV, 2013.

[19] S. Schulter, C. Leistner, and H. Bischof. "Fast and accurate image upscaling with super-resolution forests". In CVPR, 2015.

[20] R. Timofte, V. De Smet, and L. Van Gool. "A+: Adjusted anchored neighborhood regression for fast super-resolution". In ACCV, 2014.

[21] J. Kim, J. K. Lee, and K. M. Lee. "Accurate image superresolution using very deep convolutional networks". In CVPR, 2016.

[22] J. Kim, J. K. Lee, and K. M. Lee. "Deeply-recursive convolutional network for image super-resolution." In CVPR, 2016.

[23] Lai, Wei-Sheng, et al. "Deep Laplacian Pyramid Networks for Fast and Accurate Super-Resolution." arXiv preprint arXiv:1704.03915 (2017).

[24] Zeyde, R., Elad, M., Protter, M.: "On Single Image Scale-Up Using Sparse- Representations. Curves and Surfaces" (2012) 711-730

[25] Timofte, R., De Smet, V., Van Gool, L.: "Anchored Neighborhood Regression for Fast Example-Based Super Resolution". ICCV (2013) 1920-1927

[26] C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. "Patchmatch: A randomized correspondence algorithm for structural image editing". ACM Trans. on Graphics, 28(3):24, 2009.

[27] A. Singh and N. Ahuja. "Sub-band energy constraints for selfsimilarity based super-resolution". In ICPR, 2014.

[28] A. Criminisi and J. Shotton. "Decision Forests for Computer Vision and Medical Image Analysis". Springer, 2013.

[29] K. Simonyan and A. Zisserman. "Very deep convolutional networks for large-scale image recognition". In ICLR, 2015.

[30] Zeyde, Roman, Michael Elad, and Matan Protter. "On single image scale-up using sparse-representations." International conference on curves and surfaces. Springer, Berlin, Heidelberg, 2010.

[31] Bevilacqua, Marco, et al. "Low-complexity single-image super-resolution based on nonnegative neighbor embedding." (2012): 135-1.

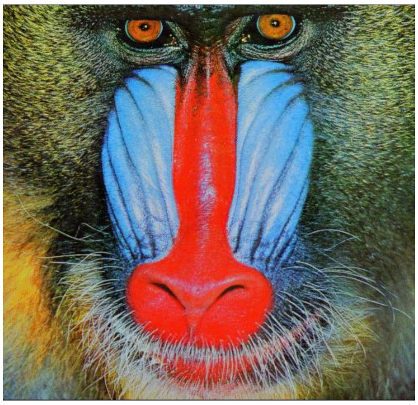
[32] Matsui, Yusuke, et al. "Sketch-based manga retrieval using manga109 dataset." Multimedia Tools and Applications (2016): 1-28.

[33] Martin, David, et al. "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics." Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on. Vol. 2. IEEE, 2001.

APPENDIX : DATASET SAMPLES:

Set 5:

Set 14:



BSDS:

URBAN:

MANGA: