



IMAGE RESOLUTION ENHANCEMENT SYSTEM USING DEEP CONVOLUTIONAL NEURAL NETWORK FOR IMAGE VISION APPLICATIONS

ALAN ABRAHAM¹, BONY BABY², AMAL MANOHAR N³, SREENIMOL K.R⁴

¹Student, Mangalam College of Engineering, Kerala, INDIA, bonybaby89@gmail.com

²Student, Mangalam College of Engineering, Kerala, INDIA, alanabraham006@gmail.com

³Student, Mangalam College of Engineering, Kerala, INDIA, amalmanohar9@gmail.com

⁴Associate Professor, Mangalam College of Engineering, Kerala, INDIA, sreenimol.kr@mangalam.in

ABSTRACT

Highlight Single image super-resolution (SR), which aims at recovering a high-resolution image from a single low-resolution image, is a classical problem in computer vision. This problem is inherently ill-posed since a multiplicity of solutions exists for any given low-resolution pixel. We propose a deep learning method for single image super-resolution (SR). Our method directly learns an end-to-end mapping between the low/high-resolution images. The mapping is represented as a deep convolution neural network (CNN) that takes the low-resolution image as the input and outputs the high-resolution one. We further show that traditional sparse-coding-based SR methods can also be viewed as a deep convolution network. But unlike traditional methods that handle each component separately, our method jointly optimizes all layers. This deep CNN has a lightweight structure, yet demonstrates state-of-the-art restoration quality, and achieves fast speed for practical on-line usage. We explore different network structures and parameter settings to achieve trade-offs between performance and speed. Moreover, we extend our network to cope with three colour channels simultaneously, and show better overall reconstruction quality. This system can be implemented in camera monitoring systems; It can be effectively used to convert low resolution images produced by systems to high quality image with with all the details.

Key words: sparse coding, CNN super resolution, convolution neural network, convolution neural network mean square error

1. INTRODUCTION

Highlight Single image super-resolution (SR), which aims at recovering a high-resolution image from a single low-resolution image, is a classical problem in computer vision. This problem is inherently ill-posed since a multiplicity of solutions exists for any given low-resolution

pixel. In other words, it is an underdetermined inverse problem, of which solution is not unique. Such a problem is typically mitigated by constraining the solution space by strong prior information. To learn the prior, recent state-of-the-art methods mostly adopt the example-based strategy. These methods either exploit internal similarities of the same image, or learn mapping functions from external low- and high-resolution exemplar pairs. The external example-based methods can be formulated for generic image super-resolution, or can be designed to suit domain specific tasks, i.e., face hallucination according to the training samples provided.

2. LITERATURE SURVEY

Methods for super-resolution can be broadly classified into two families of methods:

- (i) Bicubic Interpolation and
- (ii) Using Neural Network.

The first approach is based on a simple algorithm known as Bicubic interpolation. In this method, first the raster map of input image is determined. Then new set of pixels are added in between the pixels we already have, and the new set of pixels are filled based on the pixels surrounding them. After filling these newly added pixels based on their surroundings, the whole set of pixels are grouped together to generate the raster map of new image. This method comes under Bit-map re-sampling methods and it preserves fine details better than common bi-linear algorithms[1]. This algorithm takes low-resolution image as input, then, the semi-local interpolation is executed with maximum patch size to overcome the miss-identifications of the patches of the input image. In the second iteration, map estimation has been carried out and the anti-aliased image is given as a output. In this procedure, the semi-local interpolation is done from large to small patch sizes iteratively. The output of first iteration has been given as input to the next each interpolated semi-local masked patch centered by low-resolution pixel is compared with missing pixel centered masked patch. Most relevant LR pixel is continued and unwanted is filtered out. Then map estimation is again carried out to give the high resolution image. Most applications like Photoshop makes

use of this method for up-scaling the given image but the results were not that satisfactory.

The second method is based on neural network. The Neural Network acts as a function that takes some sort of input and turns it into some sort of output [2]. But instead of writing code for the function that does the conversion operation we train the function using a bunch of examples. In simple form it can be represented as a network that takes low resolution photo as input and turns it into a high resolution photo of the same input image and for training this network we need a bunch of low resolution image and high resolution image[3][4]. In this method we feed each of the neural networks with the pair of training datas. It takes the low resolution version of an image as input and produce the next best guess at what the up-scaled version should look like based on training data sets given. In the training process of the above method, we should have both the low resolution and high-resolution of the same image. What this method mainly focuses is on making the neural network a little bit better than previous iteration in producing the final output.

After training the network, we take a comparison between the guessed output and the real output, then we perform an operation called per-pixel subtraction and check where they actually differ. After performing per-pixel subtraction for the entire image we check the deviation from the original image to generate Loss function. The main concept behind this whole training process is to minimize the loss function. If it falls to zero then it shows complete restoration quality without any loss, this shows that our neural network is perfect. Many training sets were able to minimize the loss function to very low values but none of them were able to take it to Null value. If we compare it with Bi-cubic interpolation the results obtained were slightly better but not completely satisfactory, so there is a need for a much better method that could exhibit state of the art restoration quality.

3. PROPOSED SYSTEM

We consider image transformation problems, where an input image is transformed into an output image. Based on the best recent method, what needs to be done to get sharp and realistic image is to reduce the loss function as small as possible i.e. close to null if possible. This lead to the development of a Generative adversarial network. This method is based on adversarial games for example police thief game. When police catch a thief for his first theft he learns from experience that what mistakes did by him led to his arrest so the next time he will try to avoid them thus it can be simplified as the main concept behind this is evolve and dominate. This proposed technique utilizes the same low resolution and high resolution datas used in neural network based system as training set[7][8] The patches in the input image are identified using the method of patch extraction and representation algorithms, eg.NL means. The same neural

network from NN method is preserved in this newly proposed model as generator and this system can perform all the functionalities of the NN model i.e. comparing guessed output and original output. A new network is then added to the system called discriminator it functions like a police, it takes images as input and its output is the probability that the image is the real high resolution image or the fake high resolution image produced by the system. This discriminator network is trained so perfectly so that it can always predict whether the given image is real of fake. The architecture of the system is shown in figure-1.

Now the goal of generator is to minimize the accuracy of the discriminator. So what this does is it incentivizes the generator to produce images that looks realistic and indistinguishable. One last step is that every time we train the discriminator gives feedback to generator that it detected some images as fake so it should be made a little bit better next time so that it would be harder for the discriminator to tell. The newly proposed system is capable of producing good outcomes if iterative trainings are provided effectively. Experiments were able to show that this method provided better results that Bi-cubic Interpolation method and Neural Network based on Loss function method.

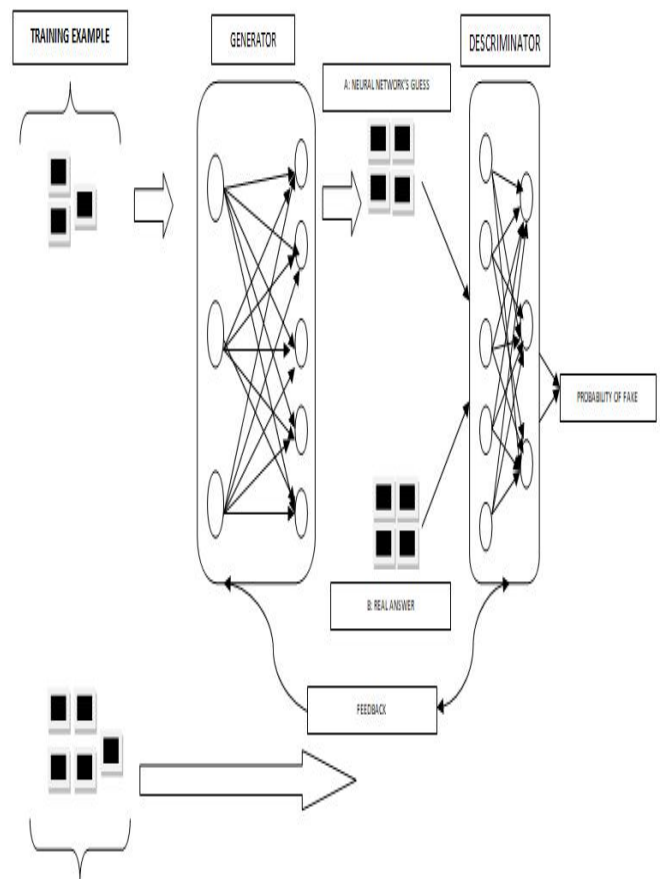


Figure 1: SYSTEM ARCHITECTURE

4. EXPERIMENTAL ANALYSIS

Single image super-resolution (SR)[1],[5],[6] which aims at recovering a high-resolution image from a single low-resolution image, is a classical problem in computer vision. This problem is inherently ill-posed since a multiplicity of solutions exists for any given low-resolution pixel. In other words, it is an underdetermined inverse problem, of which solution is not unique. Such a problem is typically mitigated by constraining the solution space by strong prior information.

The experimental analysis results showed that the proposed system was able to outperform most of the existing systems for image super-resolution. The oldest method for image super resolution is using gray world algorithm, this method was not able to generate satisfactory outcomes. Editing applications are not able to recover the details from the input image and enhance its quality but those applications are economically feasible. Applications such as Adobe Photoshop cc were able to upscale the quality of the given image but this system uses bi-cubic interpolation only. So they were not able to recover the fine details and produce state-of-the art restoration effect. Our proposed system uses a generative adversarial network that directly provides an end-to-end mapping between the low/high resolution images. This method was able to outperform both the Bi-cubic interpolation and neural network method. This method can be considered as an addition to Neural Network method that aims at reducing Loss function. The proposed method was able to mitigate the flaws in the existing model and produce satisfactory outcomes shown as in figure-3.

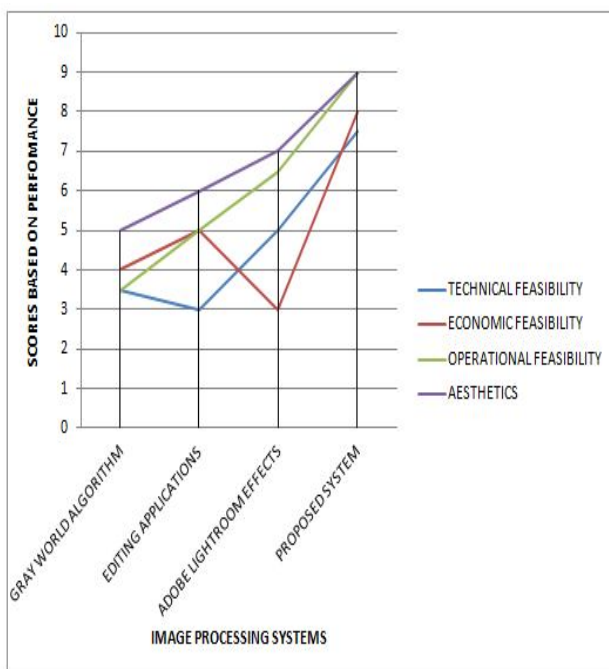


Figure 2: Experimental Analysis Graph

Considering all the flaws in the existing system, our system is able to mitigate most of these issues and produce satisfactory results. The experimental analysis shown in figure-2 indicate that our system is technically feasible, economically feasible as most of the requirements are open-source, and thus it is operational feasible. Also the system is build considering the aesthetic manner also. All these factors make the system an efficient solution for image super-resolution.

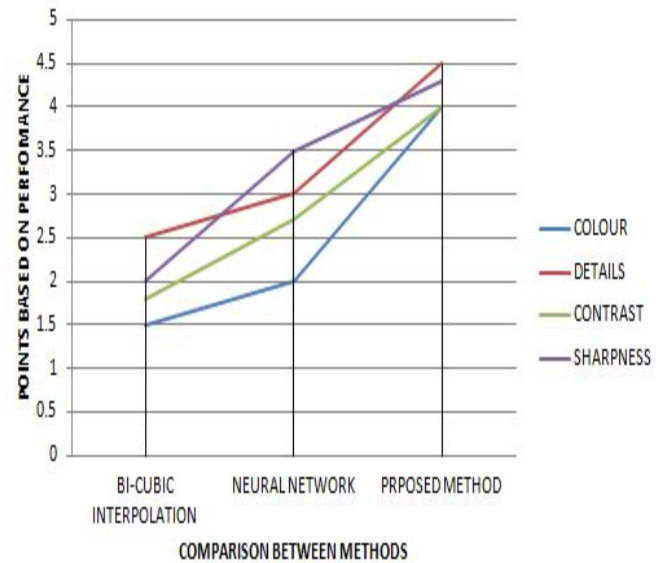


Figure 3: COMPARISON BETWEEN METHODS

5. CONCLUSION

In order to make the system more accurate we should provide large variety of training sets that consists of variety of information. The outcomes of our system is useful to solve one of the basic problem in computer vision i.e. quality image enhancement. This can be applied to recover high resolution images from low resolution images generated by CCTV cameras, identify faces and patterns, analyze images from spaces etc.

REFERENCES

- [1] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *proc. IEEE int. conf. comput. vis.*, 2009, pp. 349– 356. <https://doi.org/10.1109/ICCV.2009.5459271>
- [2] K. J. Kim and Y. Kwon, "Single-image super-resolution using sparse regression and natural image prior," *iee trans. pattern anal. mach. intell.*, vol. 32, no. 6, pp. 1127–1133, jun. 2010. <https://doi.org/10.1109/TPAMI.2010.25>
- [3] Wenzhe Shi, Jose Caballero "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network".
- [4] Jiwon Kim, Jung Kwonlee "Deeply recursive convolutional network for image super resolution".

- [5] C. Dong, C.C. Loy, K.He, and X. Tang, “Learning a deep convolutional network for image super-resolution,” in Proc. Eur. Conf. computer. vis., 2014, pp. 184–199.
https://doi.org/10.1007/978-3-319-10593-2_13
- [6] 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
<https://doi.org/10.1109/TIP.2010.2050625>
- [7] V. Jain and S. Seung, “Natural image denoising with convolutional networks,” in *proc. adv. neural inf. process. syst.*, 2008, pp. 769–776
- [8] N. Zhang, J. Donahue, R. Girshick, and T. Darrell, “Part-based renns for fine-grained category detection,” in Proc. Eur. conf. comput. vis., 2014, pp. 834–849
https://doi.org/10.1007/978-3-319-10590-1_54