Super Resolution Convolutional Neural Network For Depth Image Super-Resolution

Mr. V Veikhalou [1], Ms. Ruchika Chandel [2]
[1] Dept of Computer Science and Engineering
[2] Assistant Professor, Dept of Computer Science and Engineering
[1, 2] Alakh Prakash Goyal Shimla University, Shimla, Himachal Pradesh

Abstract- *Recently, consumer depth cameras have gained significant popularity due to their affordable cost. However, the limited resolution and quality of the depth map generated by these cameras are still problems for several applications. In this paper, we propose a new algorithm super resolution convolutional neural network. Therefore, the usage of bicubic interpolation significantly reduces the super resolution computational complexity, without sacrificing the reconstruction quality. Edge detection is one of the key stages of image processing and feature extraction. The Canny edge detector is the most popular edge detector because of its ability to detect edges in noisy images. Experimental results conducted over several images validate this result in terms of the calculated parameter SSIM and RMSE measures.*

Keywords- Low resolution by Bicubic interpolation, Iterative Back Projection, Canny edge detector, Dilated edge map, Super resoultion convolutional neural network.

I. INTRODUCTION

During recent years, we have witnessed a rapid progress in the field of 3D imaging. The birth of low-cost 3D scanning devices such as Microsoft Kinect and Time-of-Flight (TOF) cameras have opened the door for new applications in different research disciplines, including computer vision, graphics, human computer interaction, and virtual reality. However, the limited resolution and low quality of the depth map generated by these cameras still pose serious issues for various 3D applications. For example, the resolution of Swiss Range SR4000 depth camera and PMD Camcube camera are only about 200*200. Even for Kinect, the resolution of the depth image is 640*480, which is much lower compared to that of its corresponding color image (1280*1024).

In this work, we aim to enhance the resolution of depth images solely on a single depth image as input. Image super resolution (SR) aims to reconstruct a high-resolution image from its low-resolution counterpart. It is a challenging task in the computer vision field. In its essence, image SR requires the prediction of a large amount of unknown pixels based on the input pixels. To date, super resolution is also intensively related to a variety of other problems such as image denoising, deblurring, or inpainting.

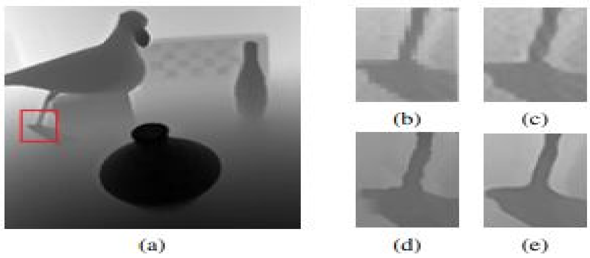

Fig. 1: Visual result (zoomed-in) on the TOF camera data up-scaled by a factor of 4.

In this paper, we address the problem of depth image super resolution and denoising, which offers unique challenges that are different from color image SR. There have been significant progresses in single image super-resolution (SR) using deep convolutional neural network. [1]

Further, this article is ordered as the follows. Section II presents various techniques & details of the proposed scheme, Section III presents literature survey of the previous scheme, Section IV presents propose work and Section V presents experiment result analysis and conclusions of the study are presented in section VII.

II. USING TECHNIQUES

In this paper we have used three techniques which are described below:

a. Low resolution by Bicubic interpolation

Image interpolation is a prime technique in image processing. It is used in many important applications such as digital high-definition television, big screen display, copy and print machine, medical imaging, end-user equipment and so on. The bicubic interpolation which is linear and easy to be computed parallel is used widely in many interpolation applications. However, the conventional bicubic interpolation has a blurring problem in the interpolated images, because it ignores the features of the image pixel data, such as the frequency features, the edge features, the features under multi-resolution and so on. To display the acquired image in full mode, we need to up-sample the image, a process that is often referred to is interpolation. One commonly used interpolation algorithm is the bicubic interpolation. It is very easy to implement in digital camera design, and is very computationally efficient. Bicubic interpolation is the default pixel interpolation algorithm. It generates each target pixel by interpolation from the nearest sixteen mapped source pixels. The interpolation artifacts such as blurring and aliasing can be greatly reduced by bicubic interpolation.

Bicubic interpolation was firstly proposed with several variations to increase the performance. The general weakness of this approach is the blur on the edge and corner of the image as the result of integrating sharpening and smoothing process. This kind of algorithm can be described as follows: Bicubic goesone step beyond bilinear by considering the closest 4x4neighborhood of known pixels--for a total of 16pixels. Since these are at various distances from the unknown pixel, closer pixels are given a higher weighting in the calculation. Bicubic produces noticeably sharper images than the previous two methods, and is perhaps the ideal combination of processing time and output quality.

Bicubic interpolation is more sophisticated and produces smoother edges than bilinear interpolation. Anew pixel is a bicubic function using 16 pixels in the nearest 4 x 4 neighborhood of the pixel in the original image. This is the method most commonly used by image editing software, printer drivers and many digital cameras for resampling images. For example, Adobe Photoshop CS offers two variants of the bicubic interpolation method: bicubic smoother and bicubic sharper. [2]

b. Canny Edge Detector

Canny edge detector uses a multi-step algorithm to detect a wide range of edges in images for edge detection. It was developed by John F Canny in 1986. Canny explaining why the technique works produced by a computational theory of edge detection. Canny operator represents the improvement of traditional single threshold method in which the high and low threshold are selected according to the gradient of the image histogram.

The primary goal of all edge detection algorithms is to locate the edge without some pre-given information. However, all algorithms are not designed to recognize high frequency impulse noise. In the Canny algorithm, Gaussian function is used to smooth the image before the edge detection process in order to reduce the Gaussian noise and to set the resolution of the image in which the change in intensity is more easily detected. All these factors contribute to precise edge detection. But canny algorithm has some weaknesses. Canny operator is not able to recognize weak edges around object. And sometime it may be possible that, it is possible to recognize false edges due to the presence of the noise. Also, it fails to recognize the edges that branch out and some important details.

Five basic steps of Canny edge detection algorithm are image smoothing and filtering, finding the gradient magnitude and gradient direction, non-maxima suppression, double threshold and edge tracking by hysteresis. The image can be smoothed by various Gaussian kernels. After the smoothing, canny algorithm finds the edges where the intensity of the gray level changes the most. These regions are located by selecting image gradients. The main aims of the Canny Edge Detector are as follows:

(a) Good detection - There should be a low probability of failing to mark real edge points, and low probability of falsely marking non edge points. So basically, we need to mark as many real edges as possible.

(b) Good localization - The points marked out as edge points by the operator should be as close as possible to the centre of the true edge. In essence, the marked out edges should be as close to the edges in the real edges as possible.

(c) Minimal response - Only one response to a certain edge. The idea is that an edge should be marked only once, and image noise should not create false edges.

The block diagram of the canny edge detection is demonstrated in Fig. 2. First smooth the image by Gaussian smoothing filter to reduce the noise contained on the image. A 5x5mask is a common choice for a size of a Gaussian filter. The next step is to calculate all pixels, direction and magnitude. By obtain vertical and parallel gradient by 3x3 Sobel operators. [3]

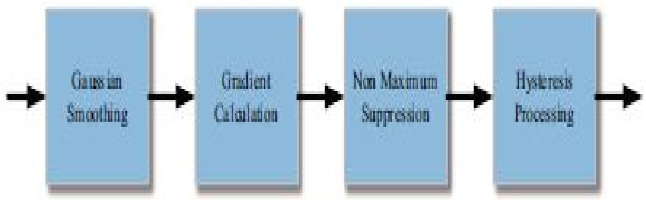

Fig. 2. Canny edge detection block diagram.

c. Dilated Edge Map

Dilated convolution consists of a contracting path to extract abstract features and an expanding path to recover spatial resolution. The features in the contracting path are concatenated to the corresponding features in the expanding path to provide the detailed image information that is lost during the successive down-sampling steps. However, the level of features in the contracting path is much lower than that in the expanding path. It will not obtain the optimal results when directly concatenating these features. To overcome this limitation, we embed the dilated dense network in the U-net to obtain a new network.

The edge map is subsequently processed morphologically by the dilation operator to extract the text regions. Dilation and erosion are the two basic operators in mathematical morphology. On a binary image, dilation gradually enlarges the boundaries of the foreground regions. Thus, areas of foreground pixels grow in size while holes within those regions become smaller. The precise effect of dilation on the input image depends on a set of coordinate points called structuring elements (or kernels). [4]

d. Super resoultion convolutional neural network

The multi-layer perceptron has been used to denoise image but all layers are fully-connected. The convolutional neural network has also been used to denoise natural image and remove noisy patterns. In the last two years, the deep convolutional neural network named SRCNN has been introduced to resolve the image super-resolution problem. The convolutional neural network can reduce the number of parameters by perceiving local area, sharing weights and sub-sampling. [5]

Recently, convolutional neural networks (CNN) have shown a tremendous outcome due to its success in image classification. Super-Resolution Convolutional Neural Network (SRCNN) not only shows the state-of-the-art performance but also does not require any features extraction that is typically used in other SR methods. In the mapping is represented as a CNN taking the LR image as the input and outputs the HR one. This work shows also a link between traditional sparse-coding-based SR methods and deep convolutional networks. SRCNN is based on a three-layers CNN: the first convolutional layer extracts a set of feature maps, the second layer maps these feature maps nonlinearly to HR patch representations and the last layer combines the predictions within a spatial neighborhood to produce an estimate of the HR image. Other deeper neural networks have been proposed for SR, demonstrating the potential of deep networks for image reconstruction and enhancement. [6]

III. LITRATURE SURVEY

Baoliang Chen et.al. [2018] In this paper, we propose single depth image super-resolution using convolutional neural networks (CNN). We adopt CNN to acquire a high-quality edge map from the input lower solution(LR) depth image. We use the high-quality edge map as the weight of the regularization term in a total variation(TV) model for super-resolution. First, we interpolate the LR depth image using bicubic interpolation and extract its low-quality edge map. Then, we get the high-quality edge map from the low-quality one using CNN. Since the CNN output often contains broken edges and holes, we refine it using the low-quality edge map. Guided by the high-quality edge map, we upsample the input LR depth image in the TV model. The edge-based guidance in TV effectively removes noise in depth while minimizing jagged artifacts and preserving sharp edges. Various experiments on the Middle bury stereo dataset and Laser Scan dataset demonstrate the superiority of the proposed method over state-of-the-arts in both qualitative and quantitative measurements. [7]

Jie Fu et.al. [2018] In this paper, a new convolutional neural-networks based super resolution(SR) is proposed. SR has been a hot research area for decades, and it includes two types: single frame-based SR and multi frame-based SR. The focus of the paper is to reconstruct the corresponding high-resolution image from a given low resolution image. The popular end-to-end learning architecture is improved and no preprocessing and image aggregation are needed. Our network model(RSCNN) uses different convolution kernels for a set of feature maps in the feature mapping step, which ensures the accuracy of reconstruction results under the premise of improving the reconstruction quality. The method is applied to Jilin-1 which is the firstself-developed commercial remote sensing satellite group in China. The results show the superiority of our method both visually and numerically by comparing with other excellent image super resolution algorithms. [8]

Jingfeng Lu et.al. [2018] Ultrasound Imaging is one of the most widely used imaging modalities for clinic diagnosis, but suffers from a low resolution due to the intrinsic physical flaws. In this paper, we present a novel unsupervised super-resolution (USSR) framework to solve the single image super-resolution (SR)problem in ultrasound images which lack of training examples. Our method utilizes the powerful nonlinear mapping ability of convolutional neural networks (CNNs), without relying on prior training or any external data. We exploit the multi-scale contextual information extracted from the test image itself to train an image-specific network at test time. We utilize several techniques to improve the convergence and accuracy, including dilated convolution and residual learning. To capture valuable internal information, dilated convolution is employed to increase the receptive field without increasing the network parameters. To speed up the convergence of the training, residual learning is used to directly learn the difference between the high-resolution and low-resolution images. Quantitative and qualitative evaluations on real ultra sound images demonstrate that the proposed method outperforms the state-of-the-art unsupervised method. [9]

Mehran Motmaen et.al. [2018] Image in painting is a restoration process which has numerous applications. Some of the in painting applications include restoring of a scanned version of an old image with scratches or removing objects in images. Different approaches have been used for implementation of inpainting algorithms. Interpolation approaches only consider one direction for this purpose. In this paper, we present a new perspective to image inpainting. We consider multiple directions and apply both one dimensional bicubic and two-dimensional hyperbolic interpolations. Neighboring pixels are selected in a hyperbolic formation for better preservation of corner pixels. We compare our work with recent inpainting approaches to show our superior results. [10]

Lata Ayesha Akter et.al. [2018] This paper proposed a new integrated image segmentation method for MRI brain images. In this method we have used a new transformation called Contour let Transform which is integrated with canny edge detector. For a better segmentation we have applied an enhancement function on the contour let coefficients before applying canny edge detector. The experimental results shows that using canny edge detector after enhancing the image by contourlet transform along with an enhancement function, the brain MRI image can be segmented very efficiently which can outperforms other conventional methods. [11]

Nguyen Viet Hung et.al. [2017] Nowadays, the developed countries have widely utilized image processing in the Intelligent Transportation System(ITS), medical, biology and other fields. The image quality enhancement is an important task, especially enhancing the resolution of the image from new interpolated pixels. The conventional Bicubic interpolation (BI) algorithm used a square matrix obtained by the original neighbor pixels, which surrounds the interpolated pixels, however the image resolution is lost in the depths and the boundary pixel value are blurred by a block artifact. This proposed algorithm aims to reduce blurred effect from block artifact by enhancing the interpolated pixel values along the edge pixels of the object of interest in the image. In other words, this paper modifies the conventional BI by adoption of Canny based edge mask for adjustment of the pixels along the detected edges in the interpolation process. The result of this proposed study can be applied to the field of camera-based traffic monitoring system. [12]

Adil Salman et.al. [2017] Plants are to be considered as one of the important things that plays a very essential role for all living beings exists on earth. But due to some unawareness and environment deterioration, some very rare plants are on the verge of extinction. Knowledge of rare leaves used for medicine and other plants is very critical in future. Leaf identification and classification plays a vital role for plant species recognition. In recent years, most of the researchers dedicate their work on leaf characterization. Leaf shape is the major parameter to classify plants. A new approach is to extract 15 features from the leaf using Canny Edge Detector and classify 22 different kinds' of plants with SVM classifier. [13]

Auangkun Rangsikunpum et.al. [2017] This paper describes a device for real-time expanding sign language images inserted in any TV programs by 2×2 scale factor. Bicubic interpolation is used as an image expansion method. The 2×2 scale factor simplifies the bicubic interpolation formula to just division by two operations which enables an efficient hardware implementation. As a prototype version, the proposed architecture is implemented on a Xilinx Zynq board with a Zynq-7020 SoC. The location and size of the expanding area can be adjusted to capture the sign language images using its on board push buttons. Experimental results show that the proposed implementation can be compatible with 1080p (1920×1080) 60 Hz HDMI video source. Target users are the deaf people who cannot clearly see the existing sign language images because of its small size. [14]

Bhagesh C. Maheshwari et.al. [2017] In today's world, we are surrounded by variety of computer vision applications e.g. medical imaging, bio-metrics, security, surveillance and robotics. Most of these applications require real time processing of a single image or sequence of images. This real time image/video processing requires high computational power and specialized hardware architecture and can't be achieved using general purpose CPUs. In this paper, a FPGA based generic canny edge detector is introduced. Edge detection is one of the basic steps in image processing, image analysis, image pattern recognition, and computer vision. We have implemented a re-sizable canny edge detector IP on programmable logic (PL)of PYNQ-Platform. The IP is integrated with HDMI input/output blocks and can process 1080p input video stream at 60 frames per second. As mentioned the canny edge detection IP is scalable with respect to frame size i.e. depending on the input frame size, the hardware architecture can be scaled up or down by changing the template parameters. The offloading of canny edge detection from PS to PL causes the CPU usage to drop from about 100%to 0%. Moreover, hardware based edge detector runs about 14times faster than the software based edge detector running onCortex-A9 ARM processor. [15]

IV. PROPOSE WORK

Problem statement:

Complex calculation compared to other two methods described above. Greater time need to generate the output compared to bilinear and nearest neighbor methods.

Propose methodology:

In this propose work we have used Contrast stretching that is a simple image enhancement in Iterative Back-Projection algorithm, process starts with the beginning guess of the HR image. Iterative Back-Projection algorithm is based on a similar thought as the computer-aided tomography where a 2-D article is reconstructed from its 1-D projections. The strategy involves a registration method, an iterative refinement for displacement estimation, and a simulation of the imaging process (the blurring impact) using a PSF. This beginning HR image can be created from the info LR image. The starting HR image is down sampled to simulate the observed LR image. The simulated LR image is subtracted from the observed LR image. In the event that the starting HR image is same as the observed HR image then the simulated LR image and observed LR image are indistinguishable and their difference is zero the HR image is estimated by back anticipating the blunder (difference) between simulated LR image by means of imaging smudge and the observed LR image. This process is rehashed iteratively to minimize the energy of the mistake. This iterative process of SR does iterations until the minimization of the cost function is attained to or some predefined iterations. Iterative Back-Projection algorithm is instinctive subsequently easy to understand. On the other hand, its ill-posed nature means that there is no novel solution.

First, we browse image from dataset and obtained this type of menu bar:

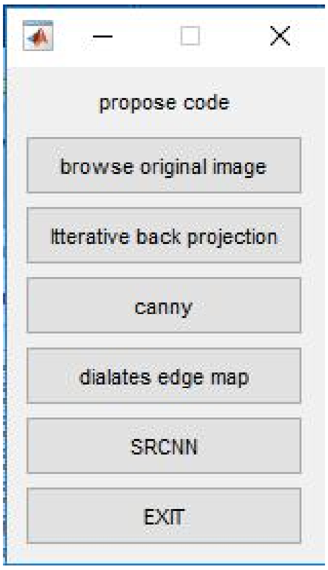

Fig.3There are 6 steps in this menu bar.

Fig.4 Browse original image from dataset.

Fig.5 Apply iterative back projection.

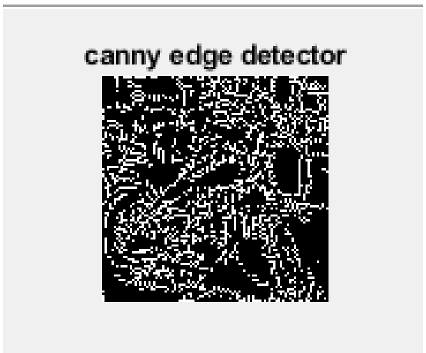

Fig. 6 Use canny edge detector.

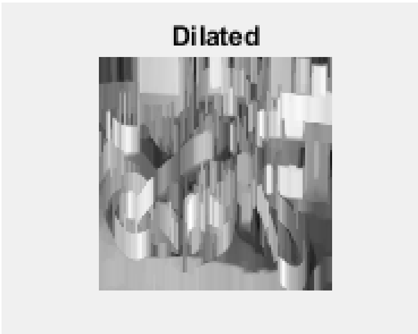

Fig.7 Dilated convolution.

Fig.8 SRCNN.

ProposeAlgorithm:

1. First, we browse image from dataset.
2. Apply iterative back projection on browse image.
3. Canny edge detector.
4. Dilated edge map.
5. SRCNN (Super resolution convolutional neural network).
6. Calculate parameter SSIM and RMSE.
7. Exit.

Flow chart:

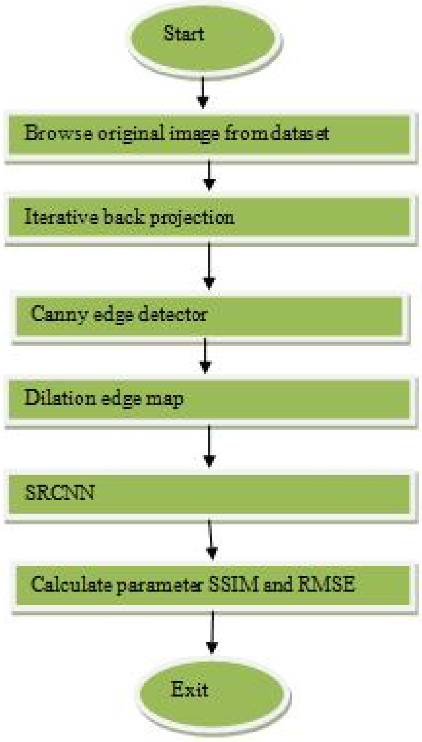

Fig 9.Flow-chart diagram of the proposed method.

Experiment Result Analysis

The picture file used in this is collected via the database. Different criteria are used to take pictures from databases and apply to create a single database. More than 20 images taken from the database include the database and everything will be converted to the similar size, data type & the resolution. The below figure shows all datasets of our experiments.

Result:

Table 1: Comparison of Base in different images.

Image name	SSIM (BI)	RMSE (BI)	SSIM (SRCNN)	RMSE(SRCNN)
Im0	0.999892	0.215653	1.056957	1.110461
2	0.999932	0.230845	0.962356	1.068152
1	0.999881	0.237858	0.991257	1.092634

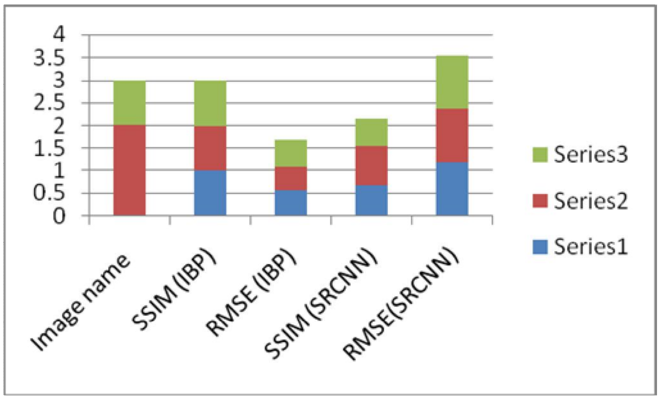

Fig 10. Graph comparison of Base in different images.

Table 2: Comparison of Propose in different images.

Image name	SSIM (IBP)	RMSE (IBP)	SSIM (SRCNN)	RMSE(SRCNN)
Im0	0.999962	0.563100	0.684312	1.196209
2	0.999961	0.527655	0.859384	1.167623
1	0.999941	0.586758	0.606912	1.175737

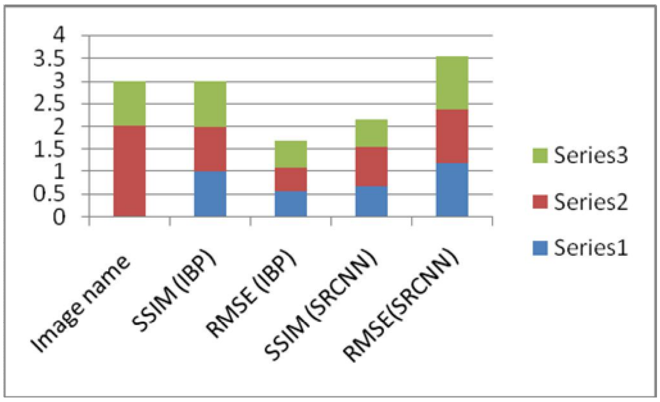

Fig 11. Graph comparison of Propose in different images.

V. CONCLUSION

In this paper, we have proposed super resolution using CNN then we used iterative back projection with canny edge detector with dilated convolution and finally apply SRCNN in our image. Besides, the performance of the extended algorithm is very similar to that of the original one. This result is validated in terms of the Calculate parameter SSIM and RMSE images over a set of test images.

REFERENCES

[1] Jun Xie, Cheng-Chuan Chou, Rogerio Feris, Ming-Ting Sun, "single depth image super resolution and denoising via coupled dictionary learning with local constraints and shock filtering", University of Washington, **Industrial Technology Research Institute (ITRI), Taiwan, ***IBM T. J. Watson Research Center Hawthorne, NY, US 2014.

[2] ZHANG Xiang-guang, "A New Kind of Super-Resolution Reconstruction Algorithm Based on the ICM and the Bicubic Interpolation" 978-0-7695-3505-0/08 $25.00 © 2008 IEEE DOI 10.1109/IITA.Workshops.2008.

[3] Akash Pandey and Dr. S.K. Shrivastava, "A Survey Paper on Calcaneus Bone Tumor Detection Using different Improved Canny Edge Detector", Computer Science and Engineering Samrat Ashok Technological Institute Vidisha (M.P.),India 2018.

[4] Hancan Zhu, Feng Shi, Li Wang, Sheng-Che Hung, Meng-Hsiang Chen, Shuai Wang, Welili Lin and Dinggang shen, "Dilated Dense U-Net for Infant Hippocampus Subfield Segmentation", Front. Neuroinform., 24 April 2019 | https://doi.org/10.3389/fninf.2019.00030 2019.

[5] Xiancai Ji, Yao Lu and Li Guo, "Image Super-Resolution With Deep Convolutional Neural Network", 978-1-5090-1192-6/16 $31.00 © 2016 IEEE DOI 10.1109/DSC.2016.

[6] Chi-Hieu Pham, Aur´elien Ducournau, Ronan Fablet and Franc¸ois Rousseau, "brain mri super-resolution using deep 3d convolutional networks", 978-1-5090-1172-8/17/$31.00 © IEEE 2017.

[7] Baoliang Chen and Cheolkon Jung, "single depth image super-resolution using convolutional neuralnetworks", 978-1-5386-4658-8/18/$31.00 © IEEE 2018.

[8] Jie Fu,Yuhong Liu and Feng Liz, "single frame super resolution with convolutional neural networkfor remote sensing imagery", 978-1-5386-7150-4/18/$31.00 © IEEE 2018.

[9] Jingfeng Lu and Wanyu Liu, "Unsupervised Super-Resolution Framework for Medical Ultrasound Images Using Dilated Convolutional Neural Networks",978-1-5386-4991-6/18/$31.00 © IEEE 2018.

[10] Mehran Motmaen, Majid Mohrekesh, Mojtaba Akbari, Nader Karimi and Shadrokh Samavi, "Image inpainting by hyperbolic selection of pixels for two-dimensional bicubic interpolations", 978-1-5386-4916-9/18/$31.00 © IEEE 2018.

[11] Lata Ayesha Akter and Goo-Rak Kwon, "Integration of Contourlet Transform and Canny Edge detector for Brain Image Segmentation", 978-1-5386-4646-5/18/$31.00 © IEEE 2018.

[12] Nguyen Viet Hung, Nguyen Thi Thu Hien, Phan Thanh Vinh, Nguyen Thi Thao and Nguyen Tien Dzung, "An Utilization of Edge Detection in a Modified Bicubic Interpolation Used for Frame Enhancement in a Camera-based Traffic Monitoring", 978-1-5386-0962-0/17/$31.00 © IEEE 2017.

[13] Adil Salman, Ashish Semwal, Upendra Bhatt and V. M Thakkar, "Leaf Classification and Identification using Canny Edge Detector and SVM Classifier",978-1-5090-4715-4/17/$31.00 © IEEE 2017.

[14] Auangkun Rangsikunpum, Ekachai Leelarasmee, Suree Pumrin, "A Design of Sign Video Image Expander for HDMI Source using Bicubic Interpolation", 978-1-5386-0449-6/17/$31.00 © IEEE 2017.

[15] Bhagesh C. Maheshwari, John Burns, Michaela Blott and Giulio Gambardella, "Implementation of a Scalable Real Time Canny Edge Detector on Programmable SOC", 978-1-5386-0872-2/17/$31.00 © IEEE 2017.