

# Design of Cascaded CNN for Medical Image Super Resolution

Kajol Singh

M Tech Scholar

Bansal Institute of Science and Technology,

Bhopal (M.P), India

kajolsingh01415@gmail.com

Manish Saxena

Professor

Bansal Institute of Science and Technology,

Bhopal (M.P), India

manish.saxena2008@gmail.com

**Abstract:** The images captured through a camera usually belong to over or under exposed conditions. The reason may be inappropriate lighting conditions or camera resolution. Hence, it is of utmost importance to have a few enhancement techniques that could make these artefacts look better. Hence, the primary objective pertaining to the adjustment and enhancement techniques is to enhance the characteristics of an image. The initial numeric values related to an image get distorted when an image is enhanced. Therefore, enhancement techniques should be designed in such a way that the image quality isn't compromised. This research work is focused on proposed a network design for deep convolution neural networks for application of super resolution techniques. To improve the complexity of existing techniques this work is intended towards network designs, different filter size and CNN architecture. The CNN model is most effective model for detection and segmentation in image. This model will improve the efficiency of medical image reconstruction from LR to HR. The proposed model showed its efficiency not only PET medical images but also on retinal database and achieved advance results as compared to existing works.

**Keywords:** Pet, DNN, HT, LBP, HR, CNN.

## I. INTRODUCTION

The past couple of years, multimedia content has generated at a rapid rate across the globe. If we look at the past, we would realize that only a very small section of the society had the privilege to possess cameras or camcorders. On the contrary, in the modern day world, digital equipment's are possessed by almost everyone we see.

Therefore, it is so easy to create the content today. As the prices went down and the features got integrated in mobile phones, these devices have become so popular. The videos and images can today be captured at the click of a button. Also the data can be stored at very cheap pricess.

A large amount of content gets created every single day. The classification, labelling and searching of content has become so easy due to its voluminous content. Almost each one of us in the modern day world uses internet for some or the other

use. Relevant content can be downloaded from any corner of the world at an instant. Classification has made the tasks simple.

The rapid increase in the amount of content has made it much more difficult to figure out people related to a particular stream of surveillance.

### A. Image

An image is generated through the arrangement of pixels in rows and columns which result into a matrix. The amplitude off is represented by selecting two spatial coordinates  $x$  and  $y$ . Using these variables, let us form a 2D function. All coordinates  $(x, y)$  in combination represent the intensity or gray level related to an image  $(x, y)$ .

### B. Processing of Digital Image

Let  $x$  represent the abscissa and  $y$  represent the ordinate which form a function  $f(x, y)$ . Any image within image processing can be represented by this two dimensional function.

Therefore, it is necessary to convert all images in form of  $f(x, y)$  function in the very first step, for which calculations has to be done. The calculations are then done by appropriate use of these values.

Hence, the process is given the name digital image processing. The reason being, all images are converted in readable format which could be easily read by a digital computer.

The intensity related to a digital image and the location as per the abscissa and ordinate are represented by pixels pertaining to the image. The locations pertaining to each pixel is different but the intensity values might be the same. Hence, in order to make proper distinction amongst pixels, it is necessary to focus on those parameters.

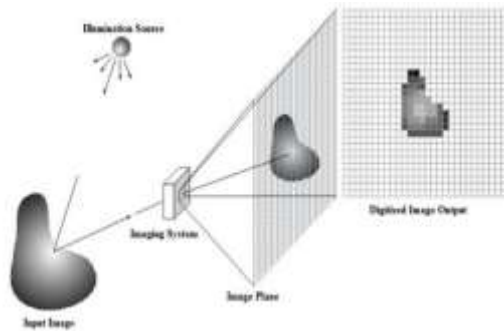


Fig. 1 Digital Image Acquisition

## II. LITERATURE REVIEW

Song et al. (2020) is used for comparative analysis purpose. In this work PET images of brain are taken as data for performance evaluation of proposed model. The author used image patches to extract the spatial location information and provided it into CNN model to identify the blur kernels in PET images. The model was designed with 20 layer of CNN with various network parameter combination.

Choi et al. (2019) proposed an architecture termed as EUSR which consists of four parts show in Discriminator and Score Predictor Multi scaling model and Multi pass Manner. Multiple scaled images from a single model is generated that's why Multi scale modeling is used.

Xunxiang, (2019) proposed a CNN model termed as SRCNN. With the pipeline of three steps that is overlapping patches, SRCNN generates high resolution output image, dense extracted and subtracting mean. For training purpose the Standard data (Set5) is used. Sparse coding is used for the deep learning and the bi cubic scaling used for the up scaling and for faster response time feed forward network is being used. One of the major advantages of this model is its faster computational quality.

George Seif et al. (2018) for the recognition purpose of image the author proposed SICNN that used 12 x 14 size image to achieve a magnification of 8x with good perceptual quality. So, for the identity metrics hyper sphere space is used, for the training data set CNN used and in the same way for the feature extraction Euclidian distance is used.

Zhu et al. (2017) performed the scaling the images with scaling factor of 2 author has used VDSR which is based on bi cubic interpolation and then for training data set with residual learning CNN is used in Zhu et al. (2017). Convergence speed is maximized and gradient clipping is used to ensure the training stability which is able to overcome the problem in SRCNN. For the multiple up scaling single model is used and it also reduces the complexity of network design.

## III. OBJECTIVES

- Improving image analysis and visualization.
- Optimization of deep network architecture.
- Improving efficiency of system.

So, this research work is focused on proposed a network design for deep convolution neural networks for application of super resolution techniques. Many researchers focused their work in designing of network learning techniques and loss function, but existing techniques are quite complex in nature which increases the overall computational complexities. To improve the complexity of existing techniques this work is intended towards network designs, different filter size and CNN architecture.

## IV. METHODOLOGY

### A. Proposed methodology

The given objectives of the thesis is performed by designing a supervised CNN model. For proposed methodology, following steps are performed:

- Step 1: Input the low contrast image
- Step 2: Generation of reference images for training of CNN model.
- Step 3: Extraction of features out of low-resolution image as well as from both reference images.
- Step 4: Training the proposed (Two layer super resolution cascaded CNN) TLSRCCNN network and generate CNN features with smallest loss function.
- Step 5: Updating the local and global gradient value according to CNN features.
- Step 6: Evaluation of Performance Parameters such as FSSIM, SSIM, MSE and PSNR values.

All the steps are discussed as below:

#### a. Input LR Image

Images from different resources are collected from different data sources or datasets that are described in details in methodology. For creating low resolution image dataset, the degradation level value is set and trained.

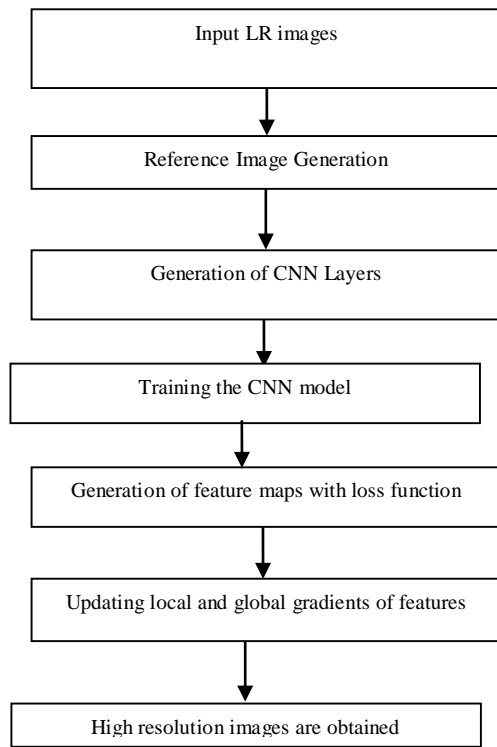


Fig. 2 Flowchart of Proposed Methodology

*b. Reference Image Generation*

For reference image generation, high resolution images are used.

➤ *Network Description*

Network Description: The robust system is developed by application of convolution neural network (CNN). In this work, residual learning is taken as proposed approach in which biblically down sampled degradation of image is mapped with residual learning. For residual learning loss function is:

$$C(\Theta) = \frac{1}{2N} \sum_{i=1}^N ||R(y_i; \Theta), y_i - x_i||_F^2 \quad (1)$$

Where,

N= number of input training images.

$y_i$  = LR image

$x_i$  = Input image

R= Residual learning function

$\Theta$  re= Input parameters for CNN.

Proposed CNN Architecture: In this work, CNN layers designed as below (Fig. 3):

- The first layer of proposed CNN architecture is combination of Conv and PReLU. In this layer 64 feature maps are taken having convolution filter of size  $3 \times 3 \times c$ .

Where,  $c$  = The number of image channel

$(c = 1)$  for grayscale or  $(c = 3)$  for RGB images.

- For non-linear function PReLU activation function is used.
- Layers 2 to 15 generates 64 feature maps, whereas from layer 16 to 29, generates feature maps increasing gradually.

In this proposed architecture, the feature maps are constructed using principal component analysis (PCA). For multivariate analysis, the PCA helps to create feature maps and for multivariate analysis. All input images are fed to PCA which generates feature maps from input images. The degradation vector  $D_v$  is fed into PCA. From each feature maps, eigen vector is calculated which is used to reduce redundant features and also extract useful information to generate reconstructed output image. The use of PCA in intermediate layers of CNN doesn't increase complexity of the overall network structure. Moreover, integration of principal components with feature maps enhance the performance.

➤ *Network Depth*

The proposed network model consists of 40 layers which takes input LR images. The LR image is created from normal image by adding degradation level. The repetitions unit of convolution and leaky ReLU activation function is taken as hidden layer and in last output layer is taken as regression layer. The output layer gives the HR whose loss function is taken and is minimized. The size of the output is set to be  $3*3*C_w$ . Where,  $C_w$  is the width of channel or feature maps. In this network the pooling layer is removed. The size of the receptive field network is for depth  $d$ .

Where,  $d = (2d + 1) \times (2d + 1)$ .

The depth selection determines the suitable trade-off between performance and efficiency level. Kernel size is another important factor that determines the efficiency and performance of the network. Generally, size of kernel is set to be  $n*n$  with degradation level of  $\sigma$ . The image size is taken to be  $w*h*c$ .

Where,  $w$ = width of image

$h$  = Height of image

$c$ = Number of channels.

Initially, in this stage, the kernel vectors are created having size  $(n*n)$  and these kernels are fed into PCA where the  $n$ -dimensional vectors are reduced into  $m$ -dimensional feature vector. Then, the low dimensional degraded image is converted into feature vector of size  $w*h*m$ . All the  $i_{th}$  map is mapped with vector  $v_i$ .

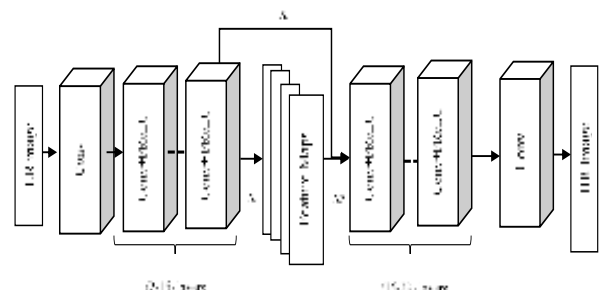


Fig. 3 Proposed Diagram of CNN Model

Then the feature maps are combined into input image and proved the effectiveness of the CNN model. This strategy will reduce the complexity of the CNN layers. Then finally cascading of these CNN layers are performed. Each CNN layer contains cascading of “Conv” and “Parametric Rectified Linear Units (PReLU)”. The last layer of the CNN is “conv” operation. The feature maps or sub images are combined and converted into image of size  $w \times h \times c$  where  $w$  is width,  $h$  is termed as height and  $c$  is termed as channel width which is 64 at initial stage and 128 after implementation of PCA. For example, LR feature vector of size  $15 \times 15$ , where  $w$  and  $h$  is 15 and channel width is 1 and converted feature maps size is  $w \times h \times (c+N+1)$  i.e.  $15 \times 15 \times (1+15+1)$ .

Table 1 Configuration of CNN Layers

Layer	Type	Strid e	Feature Maps
1	Conv	1	64
2-15	Conv+PReLU	1	64
16-29	Conv+PReLU	1	128
30	Conv	1	128

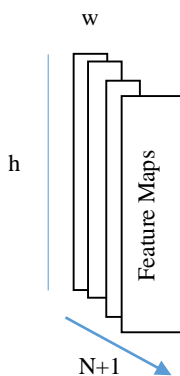


Fig. 4 Feature Map

➤ *Parametric Leaky Rectified Linear Units*

As explained in literature review, the basic CNN layers are discussed. The activation functions are used to transform image from one convolution layer to another convolution layer. The generally used activation function is rectified linear units (ReLU). The property of this function is linear in nature. This function gives positive slope only. For negative value it gives zero output. So, ReLU is considered to be non-symmetric linear function. So, it is termed as sigmoid activation functions. If input data are distributed symmetric which results in asymmetric response and increases efficiency. The ReLU activation layer is defined as below in equation 2:

$$S(x) = \frac{1}{1 + \exp(-x)} \tag{2}$$

Where,

$x$ = input data

The gradient of sigmoid activation function is defined as

$$S(x) = S(x)(1 - S(x)) \tag{3}$$

Where,  $S(x)$ = sigmoid function

When input value is large, the sigmoid function results in zero value. So, the ReLU action function is equated as below in equation (4):

$$R(x) = \begin{cases} x, & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{4}$$

When input value is less than zero, the gradient vanishes to zero. So, ReLU is not working on negative slope and train itself. In negative data, the slope in PReLU is learned and not set to zero. The random initialization of weight, the CNN with PReLU is trained. The efficiency level of PReLU is increased and enhances the efficiency and reduces the over fitting risk which occurs in ReLU and reduces the computational cost. The PReLU is defined as in equation 5:

$$P(x) = \begin{cases} x_i \\ a_i x_i \end{cases} \tag{5}$$

Where,

$a_i$  = The learnable parameter that controls the slope of negative part.

$x_i$  = The input data for the sigmoid activation function which is non-linear function  $P$  on the  $i$ th layer.

The activation function, PReLU, is used to train and learn the deep convolution neural network. This solves the gradient vanishing problem that arise in ReLU activation function. For image classification, the efficiency is improved by replacing the parameter free ReLU activation and learning problem. In this the efficiency level is improved and converges fast and reduces the complexity. The slope is parameterized in negative slope as well as positive slope. This dual parameterization controls the effectiveness of the gradient of the function. The PReLU is equated as below:

$$P(y) = \begin{cases} y_l, & \text{where } y \geq 0 \\ a_l y_l & \text{where } y < 0 \end{cases} \tag{6}$$

Where,

$a_l$  is termed as learning co-efficient or learnable parameter

$$P(y_l) = \max(0, y_l) + a_l \min(0, y_l) \tag{7}$$

If the value of  $a_l$  is set to very small, then this PReLU activation function is converted in converted into Parametric ReLU (PReLU). This reduces the zero gradient issue and leads to improvement in accuracy level. The entire network is train itself and improves the efficiency of the low-resolution images and enhances the efficiency level of the entire network. An additional small parameter is used to learn and enhances as compared to ReLU. The proposed resolution network effectively adapts and learns the PReLU parameters and jointly improves the efficiency of the entire network. A small parameter equal to small quantity and neglect the total number of weights which is equal to very small quantity.

Consequently, over fitting problem is reduced in PReLU activation function. All channels of PReLU activation function shared the coefficient  $a_l$  and in each layer. Single parameter introduced in each layer and improves the efficiency. The training is performed by back propagation and optimizes the simultaneously with other CNN layers. The experiments show that the convergence of PReLU is faster as compared to ReLU stated by Wang et al. (2020).

#### ➤ Local and Global Variable Updating

Once the extraction of feature related to a low-resolution image is done, the gradient values and intensity are modified. Hence, the design of Single Image Gradient Learner is done in such a way that weight is updated in two stages:

- The pixel intensities are updated in the first stage
- The global gradients are updated in the second stage

It is mathematically represented as:

$$I_{new}(x,y) = \sum_{n=1}^N W_n(x,y) * I_n(x,y) \quad (8)$$

$N$  = total number of images

$I_n(x,y)$  = pixel intensity related to the image

$W_n(x,y)$  = weight figured out from the features related to the input image  $I_n(x,y)$ .

$$W_n(x,y) = \exp\left(\frac{-I_n(x,y) - (1 - m_n^2)}{2\sigma_n^2}\right) \quad (9)$$

$m_n$  = the mean weight value related to  $I_n(x,y)$

$\sigma_n$  = the standard deviation of the weight value related to  $I_n(x,y)$

The global gradient is given by:

$$Wg_n(x,y) = \frac{\text{Grad}_n(I_n(x,y))^{-1}}{\sum_{n=1}^N \text{Grad}_n(I_n(x,y))^{-1} + \epsilon} \quad (10)$$

$\epsilon$  = small positive constant random value which results in non-zero denominator.

$\text{Grad}_n(I_n(x,y))$  = The cumulative histogram pixel gradient with intensity,  $I_n(x,y)$ .

As the gradient is not local to the pixels but related to all pixels that exist within a single image, it is given the name global gradient.

The LR image and their degradation is improved by applying weight updation of the CNN features. The pixels of LR images is reconstructed by enhancing the image resolution. The degraded pixels are enhanced and pixels values are improved. The degradation values are removed using updation of local as well global parametric values of the LR input image. Therefore, the proposed CNN model effectively reduces the degradation parameters and improves the effectiveness of the entire methodology.

#### B. Algorithm

Input: Input image,  $X = \{x_1, x_2, \dots, x_n\}$ , Label image,  $Y = \{y_1, y_2, \dots, y_n\}$ , degradation scale ( $\sigma$ ), Iteration (itr)

Output: Output Image ( $\hat{X}_{high\_resolution}$ )

Procedure: Image Super-resolution

Start

$X = X + \sigma$

for epochs 1:N

for iteration 1:itr

CNN  $\leftarrow X$ ; (Extracted Feature Vectors,  $FV_1$ )

CNN+PReLU  $\leftarrow FV_1$ ; (Extracted Feature)

CNN+PReLU  $\leftarrow FV_{16}$ ; (Extracted Feature)

CNN+Clip  $\leftarrow FV_{29}$ ;

end for

end for

$(\hat{X}_{high\_resolution} = Y - \text{Residue}(Y))$

Return  $\rightarrow \hat{X}_{high\_resolution}$  exit

## V. RESULTS AND DISCUSSION

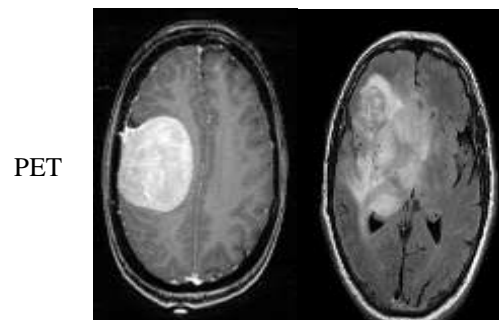
### A. Description of Dataset

In this section, on publicly we will compare the state-of-the-art algorithms in the available PET dataset [11] and retina database [12]. The emblematic images from all the datasets are shown in Figure 5.

PET images are taken from brain tumour from Kaggle repository. This contains approximate 250 images of brain tumour MRI. For simulation analysis this repository is considered. PET images are taken from brain tumour from Kaggle repository. This contains approximate 250 images of brain tumour MRI. For simulation analysis this repository is considered.

The retinal database [12] has been established by a collaborative research group to support comparative studies on automatic segmentation algorithms on retinal fundus images. The public database contains at the moment 15 images of healthy patients, 15 images of patients with diabetic retinopathy and 15 images of glaucomatous patients. The gold standard data is generated by a group of experts working in the field of retinal image analysis and clinicians from the cooperated ophthalmology clinics.

Samples of both datasets are illustrated in figure 5.



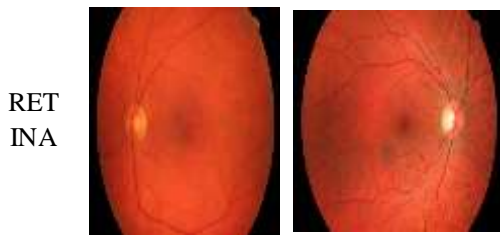


Fig. 5 Description of Dataset

**B. Performance Metric Evaluation**

To improve the quality of image, few parameters quality metrics are developed such as denoising, de blurring, color correction which varies in human visual system [HVS]. Some of Image Qualities Assessment (IQA) are PSNR, CNR and MSE. HVS are of two types: a. SSIM b. MSSIM. This help in improving quantitative and qualities measurement.

In this proposed methodology, following performance parameters are used:

*a. Peak Signal to Noise Ratio (PSNR)*

The very important tool to enhance the image quality the image is known as PSNR. The greater the value of PSNR better the technique is

$$PSNR = 20 \log_{10} \left( \frac{MAX_f}{\sqrt{MSE}} \right) \quad (11)$$

MSE is basically depends on the scaling, so the scaling algorithm gives image nearer to the original image. The best quality of image will get when there is decrease in value of mean square error.

$$MSE = \frac{1}{mn} \sum_0^{m-1} \sum_0^{n-1} ||f(i, j) - g(i, j)||^2 \quad (12)$$

*b. Structural Similarity Index (SSIM)*

One of the quality assessment methodology in digital image processing is structural similarity (SSIM). SSIM is used to find and measure of similarity between images as well as videos. This index is completely reference based and determines the noisy or compressed or distorted images with some reference images. The peak signal to noise ratio is improved with reduction of mean square error (MSE) and ultimately improved the qualitative measures.

The calculation of structural similarity formula with different window for measuring x and y of same size NxN is shown below:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (13)$$

Where,  $\mu_x$  = mean of x,  $\mu_y$  = mean of y,  $\sigma_x^2$  = variance of x,  $\sigma_y^2$  = variance of y,  $\sigma_{xy}$  = co-variance of x and y,  $c_1$  and  $c_2$  are variables to stabilize the division with weak denominator.

**C. Results and Discussions**

Table 2: Performance Evaluation of Proposed Methodology for PET

	PSNR	SSIM
	35.01408	0.965177
	37.38558	0.983774
	40.83711	0.982634
	43.56305	0.990287
	33.69888	0.952743
	39.17076	0.981139
	35.75304	0.981158
	42.67655	0.992176
	40.65057	0.98897
	35.75304	0.981158
	42.35742	0.99244
	43.45804	0.990318
	33.56344	0.95319
	36.41652	0.960185
	32.48167	0.952128
	42.17715	0.991783
	39.4019	0.981241
	51.39991	0.995056
	48.04749	0.996052
	43.53172	0.995797
	50.53024	0.997278
	47.30557	0.994023
	45.54431	0.996982
	46.39033	0.997946
	33.34988	0.956006
Average	40.81833	0.981986

Table 2 shows the performance evaluation of PET dataset on proposed model and observed average PSNR of 40 whereas average SSIM evaluated is approx. 0.9819.

Table 3: Performance Evaluation of Proposed Methodology for Retina Images

	PSNR	SSIM
	48.00665	0.988244
	48.84582	0.990822
	44.50962	0.979912
	48.10392	0.98768
	48.47755	0.989924
	43.69387	0.975607
	45.9783	0.984589
	50.95698	0.994521
	46.22785	0.984636
	46.7989	0.986377

	49.86733	0.99241
	46.26923	0.984027
	47.87547	0.98831
	49.92477	0.992526
	48.84843	0.990624
	46.23278	0.983215
	49.42559	0.991665
	47.21803	0.986543
	47.23634	0.986055
	46.59726	0.98563
	48.33531	0.989345
	45.51289	0.981342
	45.38553	0.982065
	47.76545	0.987582
	49.30531	0.990782
Average	47.49597	0.986977

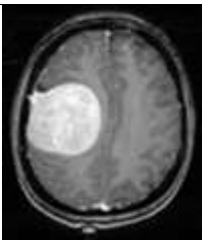



Table 3 shows the performance evaluation of PET dataset on proposed model and observed average PSNR of 40 whereas average SSIM evaluated is approx. 0.9869.

*D. Comparative Performance Evaluation*

The performance is evaluated using proposed CNN model in this section and compared with existing work. The work presented by Zhang et al. [9] is used for comparative analysis purpose. In this work PET images of brain are taken as data for performance evaluation of proposed model.

The author used image patches to extract the spatial location information and provided it into CNN model to identify the blur kernels in PET images. The model was designed with 40 layers of CNN with various network parameter combination. Table 4 represents some examples for HR resolution from LR images.

Table 4: Examples of Proposed Methodology

Dataset	LR Image	HR Image
PET		
Retina		

Below table 5 and figure 6 represent the PSNR value of proposed methodology. The table and figure show improvement of proposed methodology over existing work by 11%.

Table 5: PSNR Comparative Performance Evaluation

Techniques	PSNR
Proposed (Average)	43.5
Zhang et al. [9]	38.90

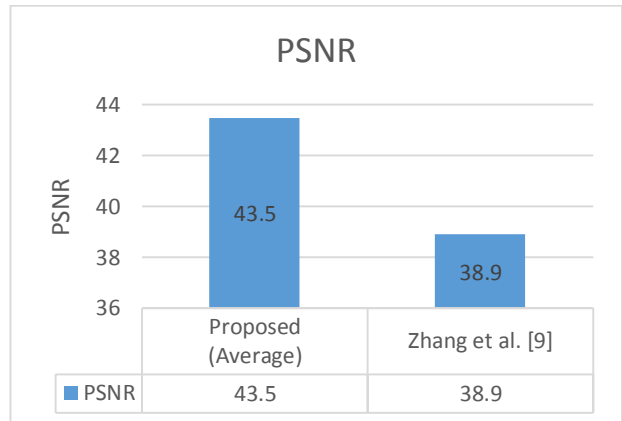


Fig. 6 PSNR Comparative Performance Evaluation

VI. CONCLUSION

In recent years, the image super-resolution (SR) has gained a lot of attention of researchers to contribute their efforts. As it is known that low resolution images with coarse details are allowed to convert corresponding image into high resolution images. This will provide a better visual effects and more refined details about image so that any application can be performed more efficiently. Image super resolution can define in terms of scaling, up sampling, enhancement, enlargement, etc. High resolution images are further applied in many applications or domains such as for object detection, contrast enhancement, medical imaging, remote sensing imaging, geo satellite images, etc. It is still a challenging task due to ill exposed images arise due to several reasons.

There are several methods or techniques that provide solution to low resolution issues. But still there are some complexities such as up scaling increases the training complexity increases. At high complexity factors, missing value recovery is very challenging task and make the entire system more complex and also assessment of quality is very low.

This research work designed a CNN model for image super resolution, termed as cascaded CNN (CAS-CNN). This model is designed with 40 layers of CNN model sandwiched with PCA feature extractor. In this model two-layer CNN model performs up scaling of LR images.

Different datasets are used for testing purposed such as retina and PET. Over all these datasets the proposed method shows improved results as compared to existing techniques. By taking both dataset, the proposed methodology shows

improved result over existing work and shows approx. 11% improvement on PSNR value.

## VII. FUTURE SCOPE

In this research work, a discussion about image super resolution, their challenges and approach to resolve such challenges. As SR is a process examines the LR images and construct high-resolution (HR) images. While during LR to HR conversion the frequency components get increased and degradations are removed. In many applications of digital world are using high resolution images for analysis. So, it arises a scope for future developments. But still, image super resolution is a challenging research issue with significant genuine applications. The achievement of any learning approaches has come about in quick development in profound convolution arrange based systems for picture super-goals. A differing set of approaches have been proposed with energizing developments in organize structures and learning systems. In future work, this model will be explored for unsupervised super resolution.

Elli Angelopoulou, „Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database,“ IET Image Processing, Volume 7, Issue 4, June 2013, pp.373-383.

## REFERENCES

- [1] Tzu-An Song, Samadrita Roy Chowdhury, Fan Yang, and Joyita Dutta. 2020. Super-Resolution PET Imaging Using Convolutional Neural Networks. *IEEE Transactions on Computational Imaging*, 6, pp. 518-528.
- [2] Jun-Ho Choi et al. 2019. Deep Learning-based Image Super-Resolution Considering Quantitative and Perceptual Quality. *Neuro Computing Elsevier*.
- [3] Xunxiang Yao, QiangWu, Peng Zhang. 2019. Adaptive rational fractal interpolation function for image super-resolution via local fractal analysis. *Image and Vision Computing*82, pp. 39-49.
- [4] George Seif et al., 2018. Edge-based loss function for single image super-resolution. *IEEE. International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. pp. 1468-1472.
- [5] Xuan Zhu. 2017. Image Super-Resolution Based on Sparse Representation via Direction and Edge Dictionaries. *Hindawi Mathematical Problems in engineering*.
- [6] KlemmGrm and Walter J. Scheirer. 2020. Face Hallucination Using Cascaded Super-Resolution and Identity Priors. *IEEE Transactions On Image Processing*, 29.
- [7] Wang, C., Li, Z., Wu, J., Fan, H., Xiao, G., Zhang, H.: ‘Deep residual haze network for image dehazing and deraining’ *IEEE Access*, 2020, 8, (1), pp. 9488–9500.
- [8] Y. Fu, T. Zhang, Y. Zheng, D. Zhang, and H. Huang. 2019. Hyperspectral image super-resolution with optimized rgb guidance, in *CVPR*.
- [9] Zhang Y, Tian Y, Kong Y, Zhong B, Fu Y. 2018. Residual dense network for image super-resolution. *IEEE conference on computer vision and pattern recognition (CVPR)*.
- [10] Ahn N., Kang B., Sohn KA. 2018. Fast, Accurate, and Lightweight Super-Resolution with Cascading Residual Network. In: Ferrari V., Hebert M., Sminchisescu C., Weiss Y. (eds) *Computer Vision – ECCV. Lecture Notes in Computer Science*, vol 11214. Springer, Cham.
- [11] <https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection>
- [12] Jan Odstrcilik, Radim Kolar, Attila Budai, Joachim Hornegger, Jiri Jan, Jiri Gazarek, Tomas Kubena, Pavel Cernosek, Ondrej Svoboda,