Proceedings of the International Conference on Augmented Intelligence and Sustainable Systems (ICAISS-2022)
IEEE Xplore Part Number: CFP22CB2-ART; ISBN: 978-1-6654-8962-1

MRI Super-Resolution using Generative Adversarial Network and Discrete Wavelet Transform

Ashwin Balasubramanian
Department of Computer Science
Loyola-ICAM College of Engineering and Technology
Chennai, India
ashwinblaze111@gmail.com

Haripriya Dhanasekaran
Department of Computer Science
Loyola-ICAM College of Engineering and Technology
Chennai, India
haripriya.23cs@licet.ac.in

Booma Raghu
Department of Electronics and Communication Engineering
Loyola-ICAM College of Engineering and Technology
Chennai, India
booma.22ece@licet.ac.in

Kunaraj Kumarasamy
Department of Electronics and Communication Engineering
Loyola-ICAM College of Engineering and Technology
Chennai, India
k.kunaraj@gmail.com

2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS) | 978-1-6654-8962-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICAISS55157.2022.10010995

Abstract— **Deep Learning Approaches have brought in major advances in super-resolution. Deep Learning architectures like convolutional neural networks or auto-encoders are trained to beget high-resolution (HR) images from the low-resolution (LR) images. Discrete wavelet transformations (DWT) is used to bring out the high frequency (HF) sub-bands of the image, which is further used to reduce the error while synthesizing super-resolution images. The proposed architecture combines the features generated by the DWT and Super-Resolution Generative Adversarial Networks (SRGAN) to produce images of high resolution. The implemented topology initially applies DWT to the image and separates out the image into two fundamental bands (HF and LF sub-bands). The features are then generated by separate generators and then Combined by inverse 2d discrete wavelet transformation (IDWT) which is then passed on to the discriminator to evaluate the generated image. A key feature of this design is that each feature is given importance and is generated by separate generators which makes the reconstruction and improvising the quality of the image more feasible. From the experiments conducted on medical images such as magnetic resonance imaging (MRI), the proposed design is computationally simpler and yet produces competitive and often improved results than state-of-the-art alternatives in terms of peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM).**

Keywords— Biomedical image processing; Image Super-resolution; Generative adversarial network

I. INTRODUCTION

Image Super-resolution reconstruction is an image processing technique that involves reconstruction of High Resolution Images from the raw Low Resolution images. Resolution of an image has always played an important role in many image processing and video-processing applications such as feature extraction, medical imaging and satellite image resolution enhancement. In the field of medicine, MRI scans that are acquired for diagnostic purposes are not of desired resolution for analysis. To address this problem wavelet techniques were introduced to reduce the noise in MRI images. Wavelets tend to play an important role in image-processing. Applications such as image construction and enhancement of low resolution images are places where wavelets are widely used. In the process of increasing the resolution of images, the clarity of images is left behind so there is an upsurge in the development of learning-based deep learning super-resolution algorithms.
Generative models have a proven performance in several applications [1, 2]. Existing methods include Stacked GAN [3] SRGAN [4] and SRCNN [5] that restore the texture details, however the images reconstructed are not robust. Here, a new GAN based model is employed to estimate HR image and alleviates the above-mentioned difficulties. In the proposed framework, wavelet transform is combined with SRGAN to reconstruct images with richer frequency details and local texture details [6].

1.1 Literature Review

A large number of studies on deep learning to improve the resolution of medical images have been carried out in the past. Significant work has been carried out and several contributions were made in the resolution enhancement of medical images using GAN, wavelet and wavelet-based GAN. A super-resolution algorithm was implemented on medical images based on several GAN versions, to obtain fruitful images and image details while eliminating the fictitious data in HF [7]. The existing excitation algorithm is enhanced by bolstering the essential features when debilitating the trivial ones. Then, by embedding the enhanced excitation algorithm in a simplified model named EDSR, a neoteric image SR block was built.

A super-resolution reformation algorithm combining DWT and GAN was proposed by Qi Zhang et al to preserve both the global information and local texture details [8]. DWT is used to enhance the high-frequency image details while GAN helps in image resolution enhancement. Deep Wavelet Super-Resolution uses all the four LR sub-bands and residuals to generate SR coefficients which is then converted into SR image [9].

Kim et al presented a highly accurate single-image SR technique that used deep CNN similar to VGG-Net and SR accuracy and visual quality was improved by augmenting the depth of the network to twenty weight layers [10]. An end-to-end mapping between the Low Resolution and High Resolution images was carried out using a light weight deep CNN after iterations of learning and this network primarily focused on real-time on-line usage [11]. A very deep CNN with 52 convolutional layers was trained effectively to map non-linearly, the LR input image with the HR target image using both residual learning [12] and recursive learning methods to optimize model parameters [13]. The SRGAN presented by C. Ledig et al is a prominent framework for synthesizing photo-realistic images [14]. This SRGAN used both content loss and adversarial loss to learn perceptual similarity and the differences in the pixels between LR and original HR images respectively. Also, interpolation techniques used in medical images are as prominent as SR techniques and a detailed comparison of the wide-spread algorithms like nearest neighbor, Gaussian interpolation truncated and windowed sine, Lagrange etc. are available. DWT coefficients were also optimized using genetic [15] programming for edge and non-edge portion of the images for improving the spatial resolution [16] and there are other significant works in implementing wavelet for interpolation and super-resolution [17, 18, 19]

1.2 Contribution

This paper uses single-level 2D-DWT of images by computing 1D-DWT along rows initially. The four sub-band coefficients encompass the entire spectrum of the actual image. Edge detection algorithm is used to identify edges in LF sub-band and the calculated edges are further used in estimating HF edges. This way the generator network learns the HF details and LF approximations of the images to generate HR images with effective storage of the texture details and improved clarity.

II. METHODOLOGY

The proposed architecture intent is to synthesize a HR image from a given input Low Resolution input image. Figure 1 shows the fundamental method for generating a synthetic High Resolution image. A random LR image taken by the generator is used to generate a synthetic HR image, while the discriminator compares it with a real HR image. The output results of the classifier are fed back to both the generator as well as the discriminator for the purpose of learning. The learning ceases to an end if the generator generates a HR image close to the real one.

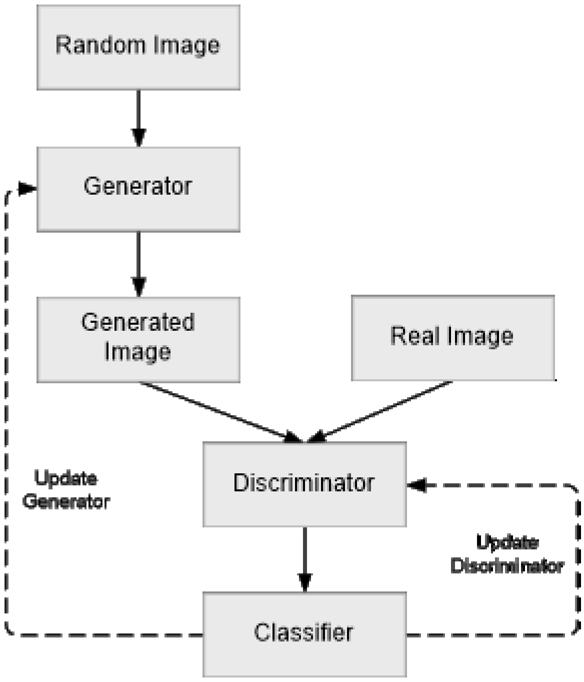


Fig. 1. Basic GAN Model for synthetic HR image

1.

2.1 PROPOSED ARCHITECTURE

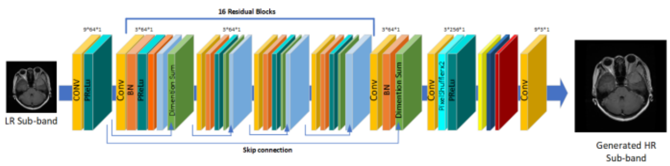


Fig. 2. Generator of each sub-band

2.1.1. Generator

The generator (figure 2) used in this architecture focuses on extracting the features of the sub-band (mainly edges) and generating an improvised image. 16 Residual Blocks are used for training purposes. The architecture resembles the one which is used by VGG16. The loss of each generator model is propagated along with the Wavelet Loss to the discriminator (figure 3) for training.

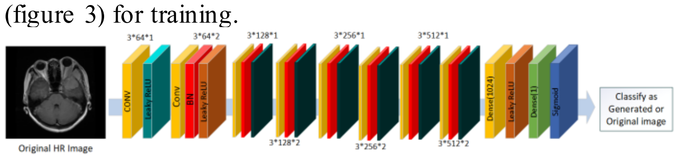


Fig. 3. Discriminator of each sub-band

2.1.2 Discriminator

The discriminator (figure 3) used in the proposed architecture mainly focuses on distinguishing between a low resolution image as well as a high resolution image. The discriminator consists of eight convolution block where the numbers of kernels are aggregated by a factor of 2 from 64 to 512 kernels.

The architecture could be viewed as a combination of SRGAN and 2D Wavelet transforms. DWT is applied to the low resolution image, for each sub-band respective generators are used sub-bands and are trained accordingly. The generated sub-bands are combined by applying inverse transformation and given as input to the discriminator. Perceptual loss specific to the DWT-SRGAN architecture shown in figure 4 is developed in order to accumulate the loss of the DWT as well as the individual generators used to generate those sub-bands.

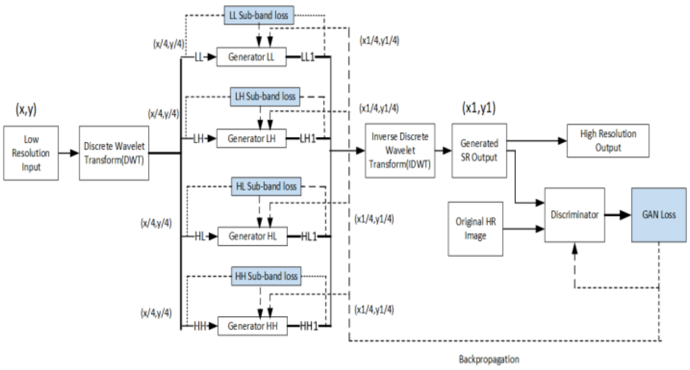


Fig.4. DWT-SRGAN architecture to synthesis HR image

III SYNTHESIS OF SR IMAGES

The raw input image is transformed by a 2D-DWT into corresponding frequency sub-bands using the bi-orthogonal wavelet. Thus, both frequency and temporal resolution can be maintained using these coefficients. Every frequency component is passed through a generator network which produces wavelet subband coefficients of increased resolution that are then transformed by inverse 2D-DWT to generate a synthetic image of higher resolution.

3.1 Perpetual Loss Function:

The perceptual loss function l^{SR} plays an integral part in sustaining the performance of the generator network. The perpetual loss of the generator of every sub-band is formulated as VGG loss, Adversarial loss and the wavelet loss.

$\text{Loss} = l_{vgg} + l_A + l_{wavelet}$

The VGG loss used here makes use of the euclidean distance between the feature representations of the generated wavelet coefficients of a low resolution image and the corresponding reference wavelet coefficient that is passed to the generator as its input. With $\Phi_{i,j}$, the feature maps are attained by the j^{th} convolution (after activation) before the i^{th} max pooling layer within the VGG network. This is the VGG loss function for every wavelet sub-band. Low-resolution wavelet sub-bands which are taken as reference images: LL, LH, HL, and HH are denoted as I^{LR}

The generated high-resolution wavelet coefficients: LL1, LH1, HL1 and HH1 are denoted as I^{HR} [20].

$$l_C = l^{SR}_{sub-band} = \frac{1}{W_{i,j}\, H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\Phi_{i,j}(I^{HR})_{x,y} - \Phi_{i,j}(G\theta_G(I^{LR})_{x,y})^2$$

Here, $W_{i,j}$ and $H_{i,j}$ describe the respective feature maps dimensions within the VGG network.

3.2 Adversarial Loss:

In supplementary to the VGG loss described so far, we also include the generative component l^{SR}_{Gen} to the perceptual loss. The generative loss is explicated based on the probabilistic values of the discriminator $D\theta_G(G\theta_G(I^{LR}))$ over all training samples.

$$l_A = l^{SR}_{Gen} = \sum_{n=1}^{N} - \log D\theta_G(G\theta_G(I^{LR}))$$

Here, $D\theta_G(G\theta_G(I^{LR}))$ is the probability that the generated image is the original HR image. For better gradient behavior, we minimize $\log D\theta_G(G\theta_G(I^{LR}))$ instead of $[1 - \log D\theta_G(G\theta_G(I^{LR}))]$.

3.3 Wavelet Loss:

The low resolution image is passed through the 2dDWT to produce four LR wavelet Sub-Bands (LRSB) which is denoted

$$LRSB = \{LRLA, LRLV, LRLH, LRLD\}$$

as: $:= 2dDWT\{LR\}$

Where the LA denotes average LR image, LH denotes horizontal LR image and LD denotes the diagonal one. 2dDWT {LR} denotes the 2dDWT of the LR image.

$$= \{HRHA - LRLA, HRHV - LRLV,$$
$$HRHH - LRLH, HRHD - LRLD\}$$

Every sub-band is passed through generator to generate High Resolution wavelet Sub-Bands (HRSB):

$$HRSB = \{HRHA, HRHV, HRHH, HRHD\}$$
$$:= 2dDWT\{HR\}$$

Where the $HRHA, HRHB, HRHH, HRHD$ represent the sub-bands of the HR image, respectively.

Then the difference of the sub bands denoted as ΔSB (residual) between corresponding LRSB and HRSB is computed as [10]:

$$\Delta SB = HRSB - LRSB$$
$$= \{HRHA - LRLA, HRHV - LRLV,$$
$$HRHH - LRLH, HRHD - LRLD\}$$
$$= \{\Delta A, \Delta V, \Delta H, \Delta D\}$$

The network learns the alterations between LR wavelet sub-bands and HR wavelet sub-bands. By summing up these differences (residual) to the input wavelet sub bands along with the discriminator loss, we get the final super resolution wavelet sub band.

The forward feeding stratagem is denoted as *f*(LRSB). The error function of the network outputs is defined as:

$\text{Cost} = \frac{1}{2}||\Delta SB - f(LRSB)||_2^2$

The weights and bias are denoted as (Θ,Ƅ).

Then the optimization problem is defined as :

$(\Theta,Ƅ) = \arg min_{(\Theta,Ƅ)} \frac{1}{2}||\Delta SB - f(LRSB)||_2^2 + \lambda \, ||\Theta||_2^2$

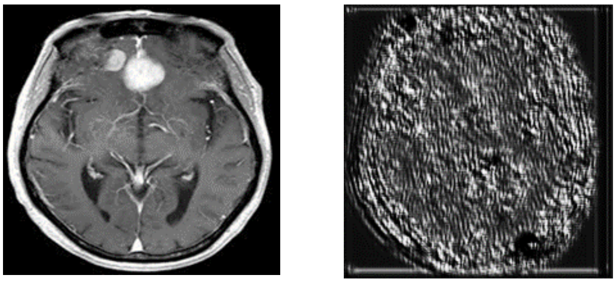
Where $||\Theta||_2^2$ denotes the weight decay regularization with parameter λ.

3.4 Training Settings

In the training process, bicubic interpolation is used to generate the LR image and then normalized with mean and standard deviation as 0 and 1 respectively. For the training purpose, MRI's of different quality and characteristics were sourced. The dataset consisted of 2500 images which are mixture of tumor Alzheimer affected and normal MRI's were included in the dataset to allow the architecture to be trained more efficiently. The dataset is transformed to HR images of size 256x256. The connection between the DWT coefficients (LR) and the generated DWT coefficients (HR) is the learning parameter for the network to generate the image. The following are the network parameters: Batch size in the network is set to 32; network learning rate = 0.0002 and the number of iterations is 100.

4. EXPERIMENTAL RESULTS

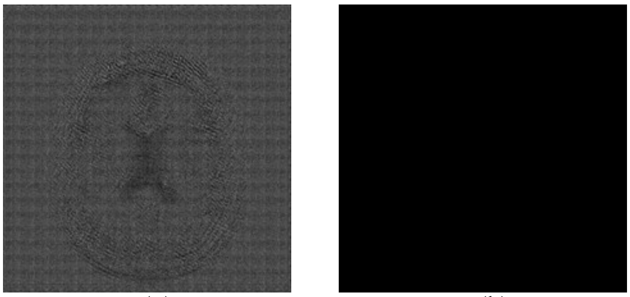
The images displayed in figure 5 are the interim resultant images that were obtained during the process of generation by the existing SRGAN algorithm. It is seen that the edges and features are not captured clearly. These were the images obtained during the training process at 25 epochs.



(a) (b)

Fig. 5. (a) Original Image (b) Output Image captured during the training process of SRGAN.

Figure 6 shows the interim resultant images generated by the proposed "Wavelet-based SRGAN" algorithm in the process of training at 10 epochs. It is observed that the presence of wavelet transform captures the edges at first which is evident from the intermediate synthetic images generated. This advantage of using wavelets transform in the model has extracted the features and texture details of the image effectively.



(a) (b)

Fig. 6. (a) Original Image (b) Output Image captured during the training process of proposed "DWT-SRGAN"

4.1 Training Plot

The training plot in the figure 7 displays the loss acquired in each iteration which is a training metric to be considered. Each iteration is the estimation of the gradient and an update of the network parameters. Also, the loss curve in figure 7 gives a snapshot of the training metric. The loss decreases as the iteration increases.

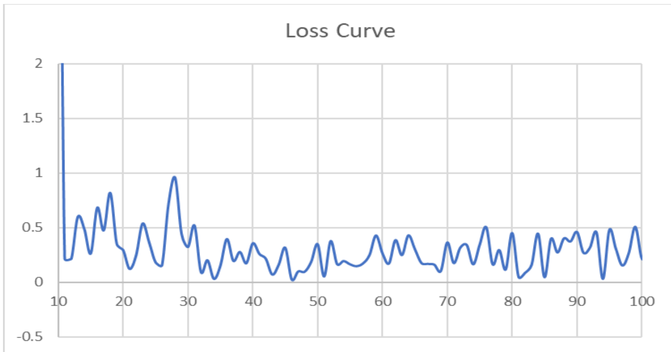


Fig.7. Loss curve of the deep neural network.
X-axis: number of iterations; Y-axis: Loss

4.2 Evaluating the trained network

The trained network is loaded into the workspace for evaluation. Sample images of MRI scans of the brain are given as input for testing the network. The model was able to generate HR images with rich frequency details and defined edges. The resultant HR images shown in figure 8 and figure 9 were generated by the model with image quality metrics,

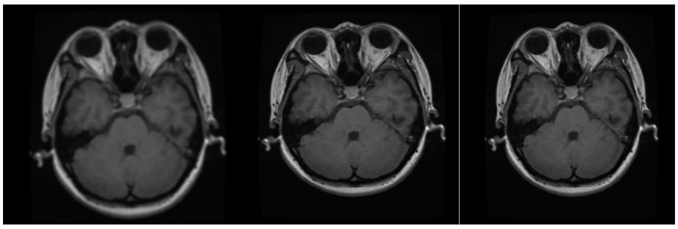
The Pixel to Noise Ratio (PSNR) and the Structural Similarity Index (SSIM) are the metrics which are used to evaluate the architecture's performance.

The PSNR can be completed by the formula

$$10\log_{10}\frac{(255)^2}{MSE}$$

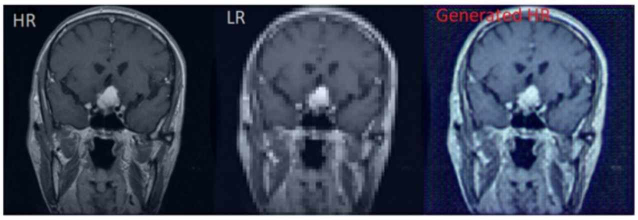
By calculating both the metrics after training the below results were generated

PSNR of 29.44 dB and SSIM of 0.92.

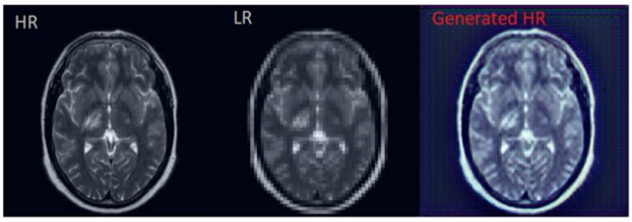


Low-resolution   Generated High-resolution   Original Image

Fig.8. Simulation Output images



9a



9b

Fig.9. Simulation interim Output images

V. Conclusion

In this paper, the authors have implemented a SR algorithm based on SRGAN and DWT to generate HR images for better visual quality particularly for medical images that help in diagnosis. The implemented DWT-SRGAN technique uses DWT to extract the HF sub-band thereby preserving the edges from the loss functions. From the experimental results, we can conclude that the average peak signal-to-noise ratio (PSNR) and SSIM is on par with the standard SR techniques available in the literature.

VI. References

[1] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," ArXiv preprint arXiv:1312.6114, 2013.

[2] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in NIPS, 2014, pp. 2672–2680.

[3] X. Huang, Y. Li, O. Poursaeed, J. Hopcroft and S. Belongie, "Stacked Generative Adversarial Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 1866-1875, doi: 10.1109/CVPR.2017.202.

[4] C. Lediget al., "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 105-114, doi: 10.1109/CVPR.2017.19.

[5] Chao Dong et al. "Image Super-Resolution Using Deep Convolutional Networks". In: CoRR abs/1501.00092 (2015). arXiv: 1501.00092. URL: http://arxiv.org/abs/1501.00092.

[6] Qi Zhang, Huafeng Wang, Tao Du, Sichen Yang, Yuehai Wang, Zhiqiang Xing, Wenle Bai, and Yang Yi. 2019. Super-resolution reconstruction algorithms based on fusion of deep learning mechanism and wavelet. In Proceedings of the 2nd International Conference on Artificial Intelligence and Pattern Recognition (AIPR '19). Association for Computing Machinery, New York, NY, USA, 102–107. https://doi.org/10.1145/3357254.3358600

[7] X. Bing, W. Zhang, L. Zheng and Y. Zhang, "Medical Image Super Resolution Using Improved Generative Adversarial Networks," in IEEE Access, vol. 7, pp. 145030-145038, 2019, doi: 10.1109/ACCESS.2019.2944862.

[8] T. Guo, H. S. Mousavi, T. H. Vu and V. Monga, "Deep Wavelet Prediction for Image Super-Resolution," 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017, pp. 1100-1109, doi: 10.1109/CVPRW.2017.148.

[9] J. Kim, J. K. Lee and K. M. Lee, "Accurate Image Super-Resolution Using Very Deep Convolutional Networks," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 1646-1654

[10] C. Dong, C. C. Loy, K. He and X. Tang, "Image Super-Resolution Using Deep Convolutional Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 2, pp. 295-307, 1 Feb. 2016. doi: 10.1109/TPAMI.2015.2439281

[11] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778. doi: 10.1109/CVPR.2016.90

[12] Y. Tai, J. Yang and X. Liu, "Image Super-Resolution via Deep Recursive Residual Network," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 2790-2798.

[13] C. Lediget al., "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 105-114.

[14] T. M. Lehmann, C. Gonner and K. Spitzer, "Survey: interpolation methods in medical image processing," in IEEE Transactions on Medical Imaging, vol. 18, no. 11, pp. 1049-1075, Nov. 1999. doi: 10.1109/42.816070

[15] D. Vaithiyanathan, R. Seshasayanan, K. Kunaraj, J. Keerthiga, "An Evolved Wavelet Library Based on Genetic Algorithm", The Scientific World Journal, vol. 2014, Article ID 494319, 17 pages, 2014. https://doi.org/10.1155/2014/494319

[16] Yuvaraj, S, Seshasayanan and Kunaraj, K, "Satellite image resolution enhancement using evolved wavelet filters," Journal of Computational and Theoretical Nanoscience, 2017, Vol.14, Issue.9, pp. 1546-1955 (14). https://doi.org/10.1166/jctn.2017.6740

[17] K. Kinebuchi, D. D. Muresan, and T. W. Parks, "Image interpolation using wavelet based hidden markov trees," in Acoustics, Speech, and Signal Processing, IEEE International Conference on, vol. 3, pp. 1957–1960, 2001.

[18] S. Zhao, H. Han, and S. Peng, "Wavelet-domain hmt based image super-resolution," in Image Processing, IEEE International Conference on, vol. 2, pp. II–953, 2003.

[19] H. Chavez-Roman and V. Ponomaryov, "Super resolution image generation using wavelet domain interpolation with edge extraction via a sparse representation," IEEE Geoscience and Remote Sensing Letters, vol. 11, no. 10, pp. 1777–1781, 2014.

[20] W. Su, Y. Qu, C. Deng, Y. Wang, F. Zheng and Z. Chen, "Enhance Generative Adversarial Networks By Wavelet Transform To Denoise Low-Dose Ct Images," 2020 IEEE International Conference on Image Processing (ICIP), 2020, pp. 350-354, doi: 10.1109/ICIP40778.2020.9190766.