Examination of Skin Cancer Images using Wavelet De-noising

B. Vasantha Lakshmi, D. Elizabath Rani

Abstract: The objective of this paper is de-noising of melanoma images using wavelets because, dermatoscopy images are corrupted by noise, which leads to fault diagnosis. Hence de-noising is essential in melanoma skin cancer image to remove the salt and pepper noise(impulse noise) by preserving the melanoma image original information. The wavelet thresholding techniques are used in this paper to de-noise the melanoma image and improved the quality of an image. Wavelet de-noising algorithm has been developed employing soft and hard thresholding techniques. It works on Daubechies, Symlet, biorthogonal wavelets at decomposition level5. Image objective performance metrics like peak signal to noise ratio, mean square error and statistical performance metrics like mean, median, standard deviation, L1 norm, L2 norm are observed and analyzed for melanoma images.

Index Terms: De-noising, mean square error (MSE), Melanoma, peak signal to noise ratio (PSNR), thresholding, wavelet transform.

I. INTRODUCTION

Early detection of skin cancer image includes Pre-processing, image Segmentation, Feature extraction, Classification. Image de-noising is an important step in Pre-processing which is the first stage in image processing. Skin cancer image enhancement improves the quality of images by removing (de-noising) the irrelevant noise present in the image and also unwanted parts in the background. For de-noising of medical images Daubechies, Symlets, and Bi-orthogonal wavelets which used with soft and hard thresholding techniques for a better image. The resultant image quality was evaluated using filter assessment parameters like Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR), and Standard deviation… Wavelet transform techniques reduced the mean square error and achieved the best value of peak signal to noise ratio (PSNR)[1-2]. Medical Image de-noising has become a vital step in the processing of medical images and removing unwanted noisy data from the image. Many image de-noising algorithms used to remove the noise and remain all the relevant features of the image. For edge preservation and multi-resolution analysis, partial differential equations are useful. The directional oriented information and texture information is very important in developing the de-noising techniques. Total variation approach and Complex Dual-Tree wavelet transform method to de-noise the medical images. The performance metrics like PSNR (Peak Signal to Noise Ratio) and SSI (Structural Similarity Index) are calculated [3-4].

B. Vasantha Lakshmi, Researcher, EECE Department, Gitam deemed to be University, Visakhapatnam.
Dr. D. Elizabeth Rani, Professor, EECE Department, Gitam University, Visakhapatnam.

Dermoscopy images are de-noised employing a mean filter, median filter, and adaptive median filters. In dermoscopy images impulse noise was most effected during image acquisition, transmission, storage and other processes which were removed with a median filter.

In Spatial filtering image intensity values (pixels) directly processed by Mean filters, Median filters, Wiener filter, etc. [5-6]. Another classification of de-noising methods, is Transform Domain Filtering techniques such as Fourier Transform, wavelet transforms. Wavelet transform is the comprehensive form of Fourier transform which signifies the kernel function by wavelets. Wavelets are definite which analyze data, based on scale and shift parameters [7-9].

The paper is organized as follows, first a qualitative assessment of HAM 10000 data set is provided. Second a brief description of wavelet decomposition and threshold selection. Third, the implementation of our proposed work and calculation of performance metrics are discussed. Fourth, the experimental results of de-noised melanoma images by brief comments and finally conclusion of our work.

II. HAM DATABASE ASSESSMENT

In this paper, HAM 10000 dataset was used for our research work. This dataset contains10016 images. This dataset was a large collection of multisource dermatoscopy images of pigmented lesions of various parts of the human body. The authors collected dermatoscopic images of various modalities with different male and female populations. In this dataset, greater than 50 percent of lesions were confirmed by pathologists from Histopathology diagnoses by scanning the ViDIR image set. While the other images using either follow-up, expert consensus or recognized by reflectance confocal microscopy. The images were extracted from a digital dermatoscopy [10]. HAM dataset was categorized based on diagnosis methods and meters as confocal, consensus, follow-up and histopathologic images.

Table-I: Analysis of HAM dataset

Image Collecting method or diagnosis	Type of skin lesion	No. of images	Total images
Confocal	Bkl	70	70
consensus	Bkl	263	902
	Df	60	
	Nv	503	
	Vasc	76	
Follow_up	Nv	3703	3703
hist	Akiec	327	5340
	Bcc	514	
	Bkl	766	
	Df	55	
	Mel	1113	
	Vasc	67	
	Nv	2498	

70 Confocal images were collected from 45 females and 25 males. This dataset has only face image with Benign keratosis (Bkl) with different age groups from 35 years to 85 years. Most of the Bkl was observed at the age of 55 years.

Table II describes skin lesions diagnosed under consensus and they were 902, out of which 422 were females and 439 males with different body parts and 41 unknown body parts of unknown persons

Table-II: Data related to consensus

Body part	Female	Male	Total
Abdomen	45	40	85
Back	73	105	178
Chest	27	13	40
Ear	2	0	2
Face	40	47	87
Foot	10	19	29
Genital	1	0	1
Hand	9	3	12
Lower extremity	90	42	132
Neck	1	10	11
Scalp	4	9	13
Trunk	41	54	95
Unknown body part	61	66	127
Upper extremity	18	31	49

Table III clearly gives the follow-up diagnosis details, a total of 3704 'nv' images including 1782 females and 1922 males with various body parts.

Table-III: Data related to Follow-up by a specialist

Body part	Female	Male	Total
Abdomen	231	339	570
Back	232	318	550
Chest	2	5	7
Ear	2	1	3
Face	9	4	13
Foot	63	54	117
Genital	27	11	38
Hand	11	8	19
Lower extremity	502	459	961
Neck	15	13	28
Scalp	2	1	3
Trunk	512	540	1052
Unknown body part	8	3	11
Upper extremity	166	167	333

In histopathology diagnosis total of 5304 images with a different type of lesions can be categorized, including akiec, bcc, bkl, df, mel, vasc, and nv.

Table-IV shows the set of data consists of 2309 females and 2976 males and 19 unknown persons with various body parts.

Table-IV: Histopathologist's information

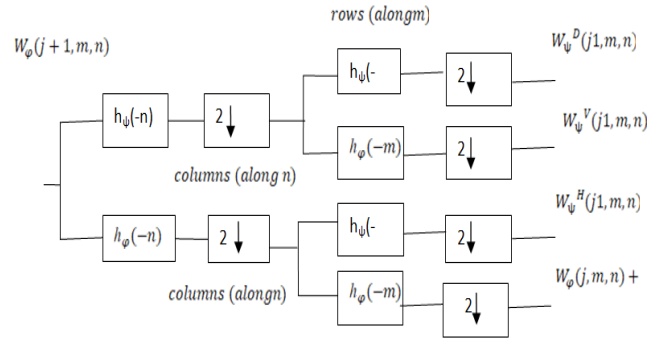| Body part	Female	Male	Unknown person	Total
Abdomen	165	166	2	333
Acral	7	0	0	7
Back	538	924	2	1464
Chest	113	247	0	360
Ear	26	25	0	51
Face	254	327	0	581
Foot	98	70	3	171
Genital	6	0	0	6
Hand	37	22	0	59
Lower extremity	560	424	0	984
Neck	57	70	0	127
Scalp	25	86	0	111
Trunk	85	169	2	256
Unknown body part	26	22	10	58
Upper extremity	312	424	0	736

Fig. 1. Wavelet transform filter bank 2-level decomposition.

III. WAVELET TRANSFORM AND THRESHOLD SELECTION

The Continuous Wavelet Transform (CWT) is the extension form of Fourier transform and it can be completely described by two equations which are transformation kernels. The Discrete Fourier transform (DWT) is a class of transformation kernel depends on the difference and also nature of the kernels. The kernels can be represented as horizontal,

vertical and diagonal components and is represented as $\psi^H(x,y) = \psi(x)\varphi(y)$; $\psi^V(x,y) = \varphi(x)\psi(y)$; $\psi^D(x,y) = \psi(x)\psi(y)$ respectively.

$\psi(x)$ and $\varphi(y)$ can be represented as linear combinations of resolution components.

$$\varphi(x) = \sum_n h_\varphi(n)\sqrt{2}\varphi(2x-n) \text{--------------------- (1)}$$
$$\psi(x) = \sum_n h_\psi(n)\sqrt{2}\varphi(2x-n) \text{------------------- (2)}$$

above equations 1 and 2 represents scaling and wavelet vectors respectively.

Fig.1 shows an iterative computational approach to the DWT. Series of filtering and downsampling by 2 operations used to compute the output coefficients like $W_\psi^D(j,m,n)$ [12]. An audio signal is a one-dimensional signal whereas an image is a two-dimensional signal, and it can be represented as x(i,j). Each row is filtered first and then down-sampled to get two images represented by (i,j/2), then every column is filtered and down-sampled to get four sub-bands named as LL, LH, HL, and HH. Hence two-level decomposition requires one 2-D scaling function and three 2-D shifting functions.

Haar wavelet is the simplest one among the wavelet family. The Haar wavelet decomposes the discrete signal into two sub-signals of equal length. One sub-signal is used to calculate the average and the other one for calculating the difference. Haar wavelet transform is given by

$$\psi_H(t) := N_1(t) = \begin{cases} 1, & \text{for } t \in [0,1/2) \\ -1, & \text{for } t \in [1/2,1) \\ 0, & \text{other wise} \end{cases}$$

$$\psi_{H,k,s}(t) = 2^{s/2}\psi_H(2^s t - k)$$

where k, s belongs to Z [11]. Daubechies wavelet has set of scaling function which is orthogonal or orthonormal. This wavelet has finite fading moments. Daubechies wavelets are more balanced in terms of frequency responses and lesser in phase responses.

Daubechies wavelets are more effective in compression and removing of noise in one- dimensional and two-dimensional signal processing because it has a property that windows are overlapped and all changes in high-frequency components reflected in the frequency coefficient spectrum.

IV. IMPLEMENTATION OF PROPOSED METHOD AND PERFORMANCE METRICS

The proposed technique is implemented on melanoma images using different types of wavelets like symlet3, bior3.7, and Daubechies on a fixed form thresholding method. Soft and hard thresholding techniques applied to achieve appreciable results. Objective performance metrics and statistical performance metrics of de-noised melanoma images were evaluated. Peak Signal to Noise Ratio (PSNR), Mean Square Error (MSE) and Variance are analyzed by the proposed algorithm. The mean square error is reduced, whereas the peak signal to noise ratio (PSNR) is achieved. Wavelet-based melanoma image de-noising procedure is explained in the algorithm given below:

➢ Image database is added to the MatLab tool box by adding a folder.

➢ Acquire the problem image from the workspace.

➢ Perform a Stationary Wavelet Decomposition. Select the desired wavelet from the Wavelet menu, select 5^{th}Level (5-level decomposition is used in this paper).

➢ If the image size is not in desired dimensions then resize the image using 'the image extension toolbox' and reload the image in workspace.

➢ Select the new image from workspace and De-noise the image using the fixed form soft thresholding method by considering soft and hard threshold values on unscaled white noise.
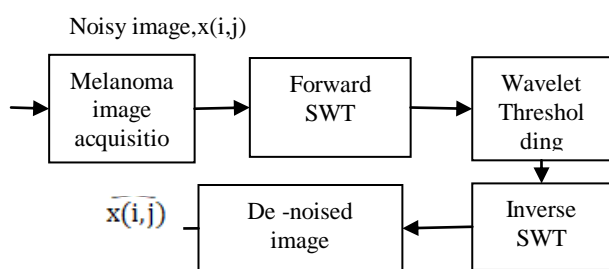
Noisy image,x(i,j)

Fig. 2. De-noising of Melanoma images

Fig.2 shows the wavelet de-noising on melanoma images. The proposed method includes image acquisition, resizing, and de-noising.

Performance Metrics: In this paper statistical performance metrics were analyzed by using bi- or3.7, Debauchees and sym3. The statistical performance metrics were mean, median, mean absolute deviation, median absolute deviation, standard deviation, L1-Norm, L2-Norm, and objective performance metrics like mean square error, peak signal to noise ratio were evaluated.

Mean is the average value of pixels and it is given by equation(3) Mean $=\sum_{i=1}^{I}\sum_{j=1}^{J} x_{i,j}$ --------------- (3)

Median gives the centre value of pixels after rearranging them in ascending order and is given as

$$\text{Median} = \frac{x_1+x_2+x_3+---+x_i+y_1+y_2+y_3+---+y_j}{I \times J} \quad -- (4)$$

Mean Absolute Deviation: This is an average distance of total elements in the data set as of the mean of the same data set, mathematically it is represented in the given equation (5).

$$\text{Mean absolute deviation} = \sum_{i=1}^{n} \frac{|\mu - x_i|}{n} \quad --------------- (5)$$

Median Absolute Deviation: It is a vigorous measure of the unpredictibility of a uni-variate sample of large data, and it is denoted as MAD, given in equation (6).

Median Absolute deviation = Median(|X-median(x)|) ----------------------------(6)

Standard deviation is a measure of contrast in the neighborhood and it was calculated using the equation (7).

$$\text{Standard deviation} = \sqrt{\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\{(y_i-x_i)(y_i-x_i)\}^2}{I \times J}} \quad ---- (7)$$

Norm: Norm is a total size or length of all vectors in a vector space. The higher the norm is larger the vector.

L1 Norm = $\sum_{i=1}^{I}(y_{i-} x_i)$ -------------------------- (8)

L2 norm: It is a performance metric that minimizing the sum of absolute differences between the actual value and the estimated value.

L2 Norm = $\sum_{i=1}^{I}\sum_{j=1}^{J}\{(y_i - x_i)(y_i - x_i)\}^2$ ----- (9)

$$\text{Mean square error} = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\{(y_i-x_i)(y_i-x_i)\}^2}{I \times J} \quad ----------------------------(10)$$

Peak signal to noise ratio = $20\log_{10}\frac{255}{MSE^2}$ ----------(11)

V. EXPERIMENTAL RESULTS

In this paper Melanoma images used for analysis are back part of different age groups of female and male persons. Experimental results of the proposed method were analyzed on the statistical performance metrics and mainly concentrated on standard deviation, mean square error, peak signal to noise ratio, mean absolute deviation and median absolute deviation. These performance metrics were analyzed by Doubecehies, Symlet and Bi-orthogonal wavelets applied on unscaled white noise.

Table-V: De-noising of melanoma images by Daubechies

Image number/parameter	Mean	Median	S D	MAD	Mean abs dev	L1 norm	L2 norm	MSE	PSNR
ISIC_0024449(soft)	0.1506	0.1348	3.957	1.969	2.941	$8.582e^{+05}$	2137	0.00732	133.5
ISIC_0024449(hard)	0.02255	0.0001399	0.6243	0.2094	0.4026	$1.17e^{+05}$	333.7	0.00115	165.7
ISIC_0024459(soft)	0.3796	0.1458	3.958	1.943	2.931	$8.555e^{+05}$	2138	0.00732	133.5
ISIC_0024459(hard)	-0.0275	0.000104	0.6077	0.1975	0.3868	$1.129e^{+05}$	328.3	0.00112	165.7
ISIC_0024516(soft)	0.5858	0.2537	3.959	1.782	2.894	$8.447e^{+05}$	2139	0.00732	133.5
ISIC_0024516(hard)	-0.0099	0	0.5782	0.1901	0.365	$1.1065e^{+05}$	312.3	0.00107	167
ISIC_0024679(soft)	0.9352	0.4342	3.962	1.954	2.974	$8.68e^{+05}$	2140	0.00733	133.5
ISIC_0024679(hard)	0.05977	-0.000564	0.6136	0.2041	0.3932	$1.148e^{+05}$	331.5	0.00113	165.7
ISIC_0024702(soft)	0.6702	0.3646	3.951	1.952	2.954	$8.62e^{+05}$	2134	0.00731	133.5
ISIC_0024702(hard)	0.08812	-0.000582	0.6009	0.1927	0.378	$1.1023e^{+05}$	324.6	0.00111	165.7

Table-V gives de-noising of melanoma images by db wavelet, fixed form thresholding on unscaled white noise (soft thresholding and hard thresholding) decomposition at level-5.

Table-VI: De-noising of melanoma images by Symlet

Image number/parameter	Mean	Median	S D	MAD	Mean abs dev	L1 norm	L2 norm	MSE	PSNR
ISIC_0024449(soft)	0.7112	0.1112	4.021	2.126	3.011	$8.787e^{+05}$	2172	0.00744	133.3
ISIC_0024449(hard)	0.03423	-0.001076	0.569	0.2815	0.4072	$1.188e^{+05}$	307.4	0.00105	167.2
ISIC_0024459(soft)	0.189	0.1465	4.019	2.134	3.012	$8.791e^{+05}$	2171	0.00743	133.3
ISIC_0024459(hard)	0.0667	-0.000444	0.5669	0.2786	0.4045	$1.18e^{+05}$	306.2	0.00104	167.2
ISIC_0024516(soft)	0.8326	0.2723	4.023	2.046	2.997	$8.748e^{+05}$	2173	0.00744	133.3
ISIC_0024516(hard)	0.01918	-0.001263	0.5713	0.2844	0.4094	$1.195e^{+05}$	308.7	0.00106	167.2
ISIC_0024679(soft)	0.729	0.4703	4.023	2.093	3.041	$8.87e^{+05}$	2173	0.00744	133.3
ISIC_0024679(hard)	0.02361	-0.000316	0.5662	0.2805	0.405	$1.182e^{+05}$	305.8	0.00104	167.2
ISIC_0024702(soft)	1.005	0.3485	4.022	2.172	3.048	$8.894e^{+05}$	2173	0.00744	133.3
ISIC_0024702(hard)	-0.04439	-0.000973	0.5593	0.2743	0.3986	$1.163e^{+05}$	302.1	0.00103	167.2

Table-VI gives de-noising of melanoma images by sym3 wavelet, fixed form thresholding on unscaled white noise (hard thresholding) decomposition at level-5.

Table-VII: De-noising of melanoma images by bior3.7

Image number/parameter	Mean	Median	S D	MAD	Mean abs dev	L1 norm	L2 norm	MSE	PSNR
ISIC_0024449(soft)	0.1895	0.05586	4.428	2.876	3.509	1.024^{+06}	2392	0.00819	131.6
ISIC_0024449(hard)	-0.07082	-0.001593	0.9088	0.4225	0.6402	$1.868e^{+05}$	491	0.00168	159
ISIC_0024459(soft)	0.4954	0.08562	4.429	2.882	3.513	1.025^{+06}	2393	0.00819	131.6
ISIC_0024459(hard)	0.2587	0.0002866	0.9121	0.4214	0.641	$1.871e^{+05}$	492.7	0.00169	159
ISIC_0024516(soft)	0.5586	0.1793	4.423	2.852	3.502	1.022^{+06}	2390	0.00818	131.6
ISIC_0024516(hard)	-0.02877	$-6.883e^{-05}$	0.9182	0.4293	0.6479	$1.891e^{+05}$	496.1	0.00169	159
ISIC_0024679(soft)	1.127	0.3419	4.425	2.842	3.512	1.025^{+06}	2391	0.00819	131.6
ISIC_0024679(hard)	-0.02057	-0.000644	0.9111	0.4225	0.6423	$1.874e^{+05}$	492.2	0.00169	159
ISIC_0024702(soft)	0.8748	0.2305	4.426	2.894	3.521	1.027^{+06}	2391	0.00819	131.6
ISIC_0024702(hard)	0.02767	-0.000396	0.9117	0.4217	0.6412	$1.871e^{+05}$	492.5	0.00169	159

Table-VII gives de-noising of melanoma images by bior3.7 wavelet, fixed form thresholding on unscaled white noise (hard thresholding) decomposotion at level-5.

Table-VIII: Comparison of MSE and PSNR

Image number	Median filter		Wavelet Transform	
	MSE	PSNR	MSE	PSNR
ISIC_0024449	6.6368	79.8	0.00744	133.3
ISIC_0024459	5.1876	81.9	0.00743	133.3
ISIC_0024516	2.2496	89.2	0.00744	133.3
ISIC_0024679	6.3545	80.2	0.00744	133.3
ISIC_0024702	8.7785	77.4	0.00744	133.3

Table VIII shows comparison of mean square error and peak signal to noise ratio for median filter and symlet3.
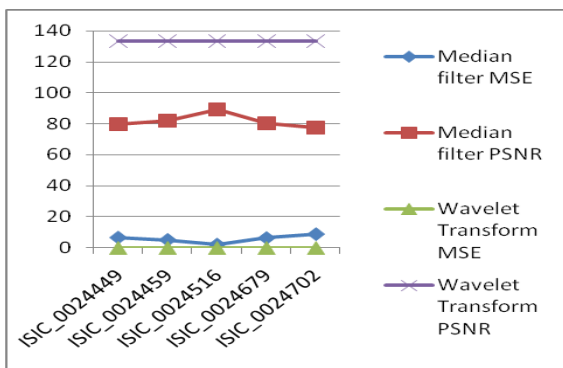
Table-IX: Comparison table for various noise densities.

Noise density	MSE	PSNR
10%	5.10	82.1
20%	6.53	79.95
30%	9.38	76.8
40%	14.28	73.16
50%	23.05	69.0
60%	35.55	65.2
70%	53.3	61.7
80%	75.6	58.6

Table-IX shows an apparent variation of the mean square error and peak signal to noise ratio in dB over an image with different impulse noise densities by median filtering of melanoma skin cancer image of ISIC_0024449.
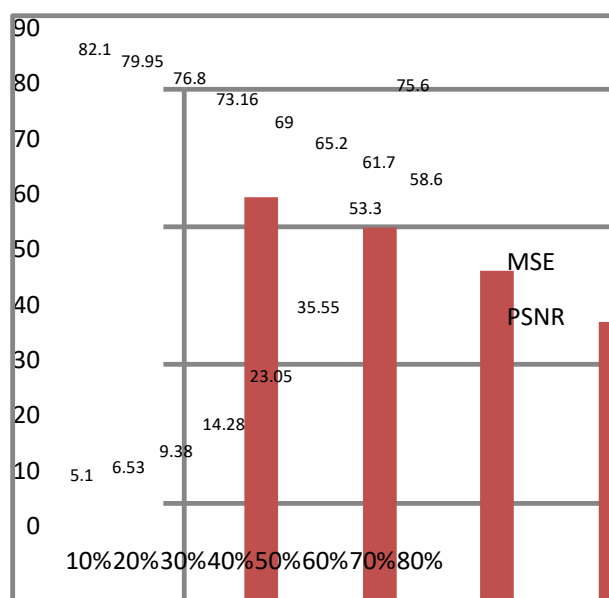

Fig.3.Comparison of the Median filter and wavelet transform output in terms of MSE and PSNR

Fig.4. Comparison of mean square error and peak signal to noise ratio with different noise densities.
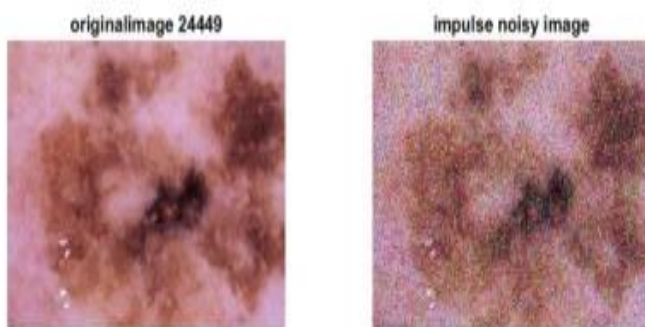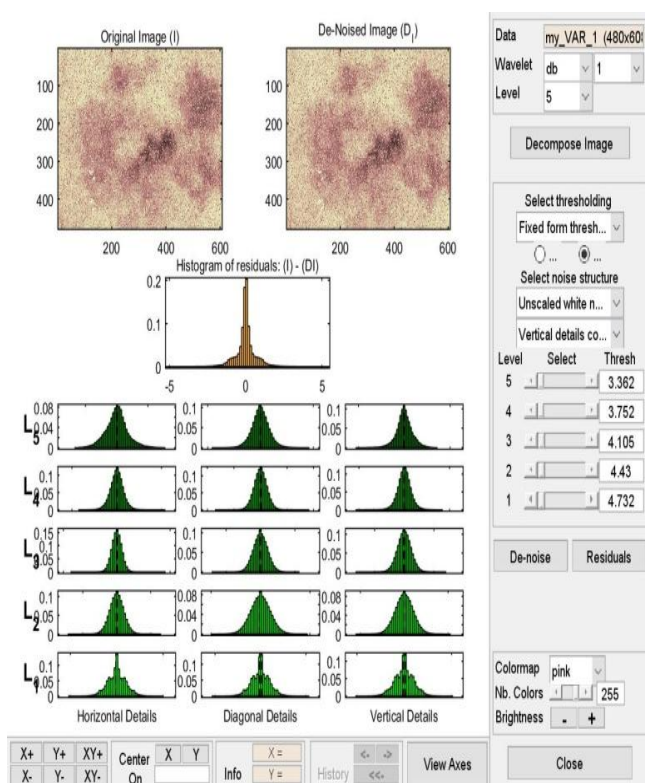

Fig.5. Sampled original image and noised image

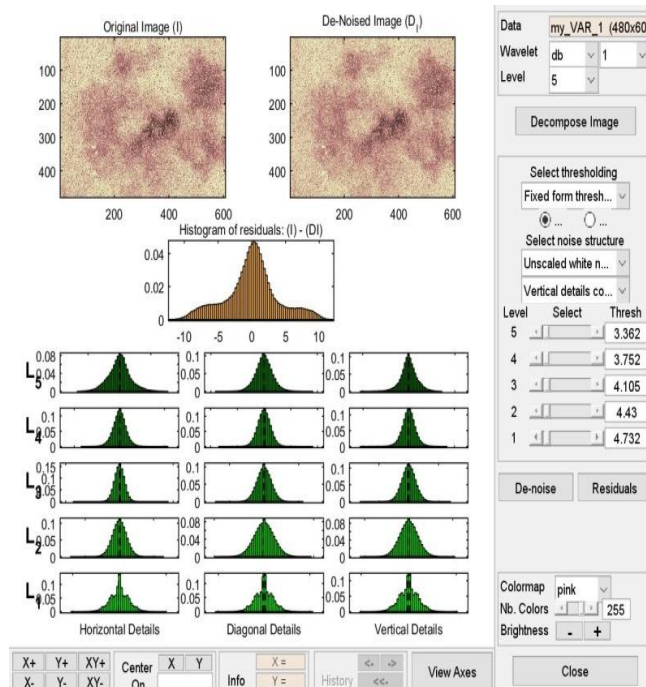**Fig.6. Image ISIC _0024449 de-noising using hard threshold**

Fig.7. Image ISIC _0024449 de-noising using soft threshold.

VI. CONCLUSION AND FUTURE WORK

De-noising is a critical task, particularly in melanoma images. In this paper, impulse noise with 20% density is added to the original melanoma image. De-noising of melanoma image is implemented by soft thresholding and hard thresholding method using, symlet3, biorthogonal3.7, and Daubechies wavelets. Symlet and Daubechies are more efficient in de-noising of melanoma image particularly in the removal of un-scaled white noise. Analysis of the results discloses that the proposed algorithm reduces the mean square error (MSE) and improves the peak signal to noise ratio. Experimental results show that Daubechies and symlet3 wavelets have similar MSE and PSNR values whereas biorthonal3.7 wavelet with smaller MSE and PSNR. We get a larger standard deviation for the soft thresholding method and a smaller value for the hard thresholding. The average distance of all pixels within the image is mean absolute deviation. It is 3 for wavelet db and is in the range of 3to3.5 for sym3 and bior3.7. Median absolute deviation(MAD) gives the variability of a univariate sample, which is less than 2 for the soft thresholding method and about to 0.2 for the hard thresholding in Daubechies. Its value is less than 0.5 in the hard thresholding and greater than 2 for the soft thresholding in bior3.7, sym3. By observing the histogram it follows the Gaussian distribution for the soft threshold, whereas non-Gaussian distribution for the hard thresholding method. Performance metrics like MSE and PSNR are calculated on 20% noise density using median filter and symlet3. The performance of Symlet3 is better than the median filter. One more important observation is one test image is subjected to various noise densities and concluded that lower the mean square error (MSE) the higher the peak signal to noise ratio(PSNR).

The de-noised images will be applied as input for image segmentation using various edge detectors as my future work.

REFERENCES

1. N.Mustafa, S.Ahmed, Jiang.P.Li, Mohaned Giess "Medical Image De-Noising Schemes using Wavelet Transform with FixedForm Thresholding", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 6, No. 10, 2015.
2. T.Grover, "Denoising of Medical Images using Wavelet Transform" Imperial Journal of Interdisciplinary Research(IJIR), Volume-2, Issue-3,2016.
3. R.E.Woods and Rafael C. Gonzalez "Digital Image Processing", Second Edition, Pearson Education.
4. Dr.T.Venkateswarlu, and V.N.PrudhviRaj "Denoising of Medical Images using Image Fusion Techniques", Signal & Image Processing: An International Journal (SIPIJ) Vol.3, No.4, August 2012.
5. A.NooriHoshyar, Afsaneh N. Hoshyar, A.Al-Jumailya, 'The Beneficial Techniques in Pre-Processing step of Skin cancer Detection system comparing' an "International Conference on Robot PRIDE 2013-2014 - Medical and Rehabilitation Robotics and Instrumentation", Conference Pride, 2013-2014
6. B.Vasantha Lakshmi, Dr. D. Elizabeth Rani, " De-noising Of Skin Cancer Images Using Wavelet Thresholding And Analysis Of Statistical Performance Metrics", Journal of Emerging Technologies and Innovative Research, Volume-6, Issue No-5, May-2019.
7. Amir Reza Sadri, MaryamZekriSaeid Sadri, NiloofarGheissarri, "Impulse Noise Cancellation of Medical Images using Wavelet Networks and Median Filters", Journal of Medical Signals and Sensors, Vol-2, Issue-1, January-March 2012.
8. Y.Yitzhaky,L.Graham, and Ibrahim Abdulhalim, "Classification of skin moles from optical spectro polarimetric images: a pilot study", Journal of Biomedical Optics 18(11), 111403, November 2013.
9. M. Srivastava, C.L.Anderson, J. H. Freed, " A New Wavelet De-noising Method For Selecting Decomposition Levels And Noise Thresholds". IEEE Access. 2016; 4:3862- 3877, doi:10.1109/ACCESS. 2016. 2587581.
10. Philipp Tschandl, Harald Kittler and Cliff Rosendahl "Data Descriptor: The HAM 10000dataset, A Large Collection Of Multi-Source Dermatoscopic Images Of Common Pigmented Skin Lesions", Scientific Data 5:180161, DOI:10.1038/sdata.2018.161
11. Ales Prachazla, Martina Mudrova "Wavelet- Based Image Segmentation", https:// www.researchgate.net/ publication/ 228453305.
12. Rafael C.Gonzalez, Steven L.Eddins, and Richard E. Woods, "Digital Image Processing Using MATLAB", Pearson Education.

AUTHORS PROFILE

B. Vasantha Lakshmi received the Engineering degree in Electronics and Communication Engineering from Jawaharlal Nehru Technological University, Kakinada in 2006, the M.Tech degree in DECS from Gudlavalleru Engineering college, University of JNTUK, Kakinada. Currently doing research in Gitam deemed to be University, Visakhapatnam.

Dr.D.Elizabeth Rani, Professor EECE Department, Gitam University, Visakhapatnam. Life Member of the Institution of electronics and Telecommunication Engineers, Life Member of society of EMC Engineers, Life Member of Society for Technical Education. She was awarded as a best teacher in the year 2017.