**International Journal of Engineering and Information Systems (IJEAIS)**
**ISSN: 2000-000X**
**Vol. 2 Issue 12, December – 2018, Pages: 18-23**

# Survey of Denoising Techniques in Image Processing

**Khattab O. Khorsheed[1], Omar G. Abood[2]**

[1] *Ministry of Education/General Directorat of Education in Kirkuk, Egypt.*
[2] *Department of Information Technology, Institute of Graduate Studies and Researches, Alexandria University, Egypt.*
*E-mail*: omar.ghazi88@yahoo.com[1], komer592@gmail.com[2]

***Abstract*****:** *The reduction of noise from a signal remains to be a problematic task for researchers. While several works have been published on algorithms, each of these methods have their rewards and their limits. This work showcases a review of some of the major works developed in this field of image noise reduction. The work starts out with an introduction then presents some of the most widely used approaches in their classification-sets and an outline of numerous algorithms and analysis is provided. Intuitions and possible future trends in the area of reducing noise are also presented herein.*

**Keyword:** *Digital Images, Wavelet Transform, Undecimated Wavelet Transform, Denoising Images.*

## 1. Introduction

The role of digital images plays a crucial role in daily life where it is used in applications like satellite TV, magnetic resonance imaging, computer tomography along with other fields of research and technology including Geographical Information Systems (GIS) alongside astronomy and other scientific fields. Data sets gathered by image sensors are normally noise-tainted. Flawed tools interfere with the data acquisition procedure and interfering natural phenomena can all decrease levels of the data of interest [1]. Another way to introduce noise is through the transmission errors and compression. It is for that reason that denoising is deemed essential at times and the initial step to be taken before the data images is investigated. It has become a necessity that applying an efficient denoising method to make up for the corruption of data at hand.

Denoising images continues to be a challenge for researchers due to the fact that removing noise brings up artifacts and causes blurring of the images. The work presented herein describes various methodologies for the reduction of noise (or denoising) providing an overview as to which algorithm should be used to determine the most reliable approximation of the initial image data provided its corrupted form.

The modelling of noise in images is majorly affected through capturing instruments, data transmission media, image quantization and discrete sources of radiation. A variety of algorithms is used in regard to the noise model. The majority of the natural images are thought to have additive unplanned noise modeled as a Gaussian. Speckle noise is seen through ultrasound images whereas Rician noise [1] disturbs MRI images. As mentioned, the work emphasizes on noise removal techniques for natural images.

## 2. Evolution of Image Denoising Research

The field of image denoising has remain to be an essential obstacle in image processing. It is shown that the performance of wavelets in image denoising is better because of their assets including sparsity and multiresolution structure. With Wavelet Transform (WT) becoming more and more popular in the past two decades, a wide variety of (WT) algorithms were produced. It was thus that the attention has moved from the Spatial and Fourier domain towards the (WT) domain. Ever since the introduction of Donoho's Wavelet based threshold approach back in 1995, a massive rush in the denoising papers were accomplished. While the concept Donoho introduced was not extraordinary, said methods needn't following or connection of the maximum and minimum for the wavelet across the different scales as suggested by Mallat [2]. This has revived the attention paid in wavelet based denoising techniques since Donoho [3] established a simple method to a complex problem. Researchers in the field have accomplished different manners to compute the parameters for the thresholding of wavelet coefficients. Data adaptive thresholds were presented to attain maximum rate of threshold. In the following periods, work found that significant developments in perceptual quality may possibly be found through translation invariant methods based on the thresholding of an UWT [5].

The thresholding techniques mentioned have been applied to the non-orthogonal wavelet coefficients to help decrease the artifacts. Multiwavelets have been used in finding alike results whereas probabilistic models that employ the statistical properties of the wavelet coefficient tended to outdo the thresholding techniques and obtained more ground. In recent times, a lot of work has been given to the Bayesian denoising in the wavelet domain. Hidden Markov Models and Gaussian Scale Mixtures have achieved renowned popularity while research continues to be published. Tree Structures organizing the wavelet coefficients depending on their scale, scale and spatial location that have been researched. Several data adaptive transforms have been studies, among those is the Independent Component Analysis (ICA). The trend remains to pay attention to using diverse statistical prototypes to model the statistical assets of the wavelet coefficients and its neighbors. Future works will be aimed to find clearer probabilistic models for the supply of non-orthogonal wavelet coefficients.

## 3. Classification of Denoising Algorithms

Figure 1 presents the two methods mainly used in image denoising: spatial filtering and transform domain filtering.

### 3.1 Spatial Filtering

What Spatial Filtering presents is an orthodox manner of removing noise from images through using spatial filters. Those could be identified along as non-linear and linear filters.

### I. Non-Linear Filters

Through employing non-linear filters, the noise could be erased without any tries to classify it. It is thus that spatial filters use a low pass filtering on a number of pixels, assuming that the noise uses up the higher frequency spectrum region. In general, spatial filters' job is to remove the noise towards an acceptable rate; however, this comes at the price of causing images to blur in what causes the edges of the pictures to be unrecognizable. A number of nonlinear median filters were developed in the past few years such as weighted median rank conditioned rank selection and relaxed median [10] have been developed to overcome this disadvantage.

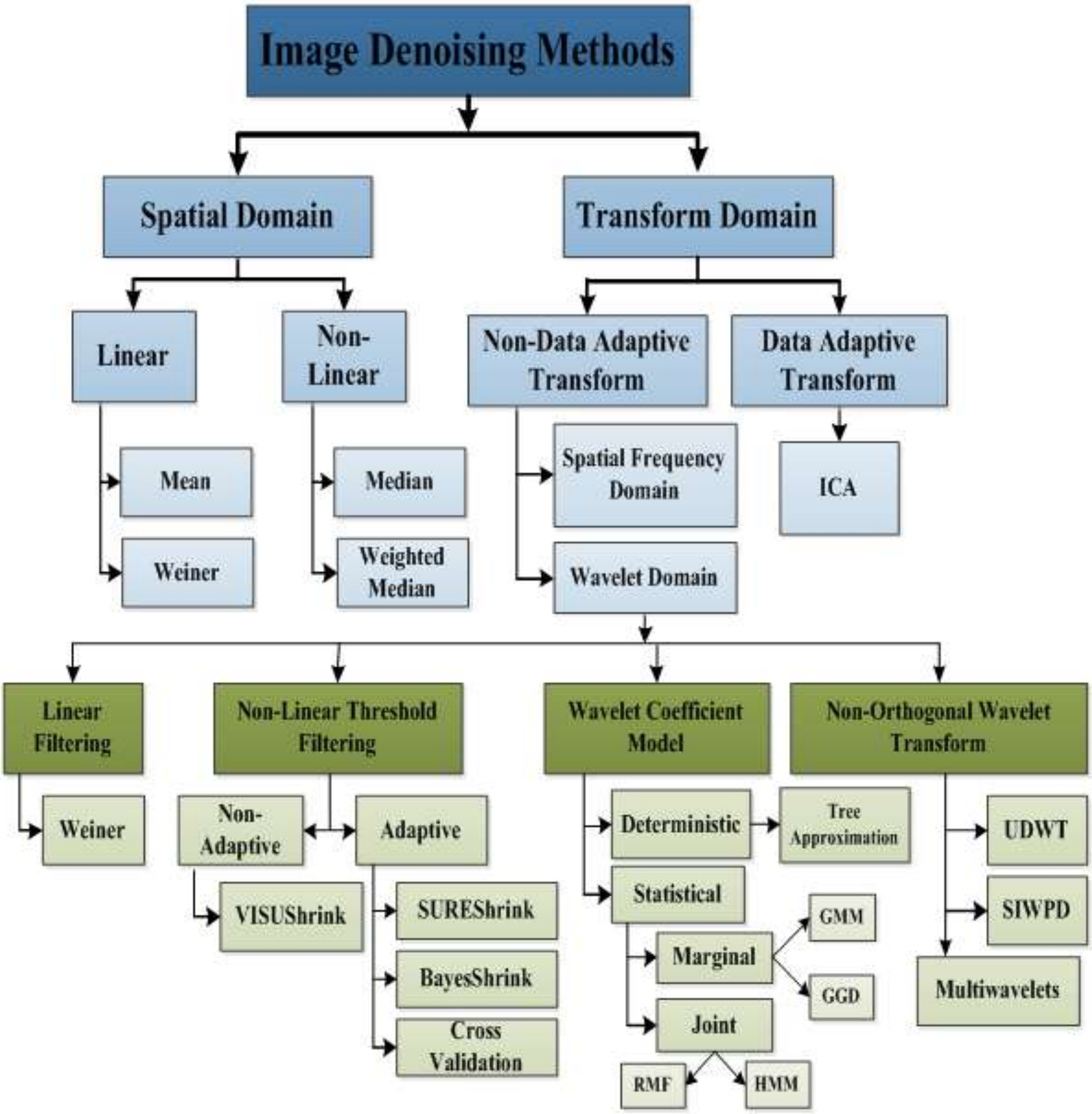


Figure 1 Classification of Image Denoising Methods

### II. Linear Filters

A mean filter is regarded as the most ideal filter for Gaussian noise in the domain of mean square error. However, it should be kept in mind that linear filters have a tendency of blurring the sharp edges which means the complete removal of some lines and other fine details. That demonstrates the weak performance in the presence of signal-dependent noise. The wiener filtering method [6] uses information regarding the spectra of noise and the original signal and it shows that it can only achieve if the fundamental signal is shown to be smooth. It employs spatial leveling and the model complexity control in order to help with choosing the size of the window and in order to beat the current weaknesses and disadvantages of the Wiener filter. For those weaknesses, Donoho and Johnstone presented a wavelet model based on the denoising scheme in [7].

### 3.2 Transform Domain Filtering

The transform domain filtering methods can be segmented in line with the choice of the basic functions. The basic purposes could be categorized as data adaptive and non-adaptive. Non-adaptive transforms are first discussed due to their surmounting popularity.

#### 3.2.1 Spatial-Frequency Filtering

Spatial-frequency filtering prefers employing low pass filters to using Fast Fourier Transform (FFT). In frequency smoothing methods [6] the elimination of the noise is obtained by creating a frequency domain filter and acclimatizing a cut-off frequency when the noise components are decorrelated from the beneficial signal in the frequency domain. These approaches are time uncontrollable and depend on the cut-off frequency and the behavior of the filter function. Additionally, they could create imitation frequencies in the processed image.

#### 3.2.2 Wavelet domain

Within the wavelet domain, the filtering operations are divided into two methods: linear and nonlinear.

**I. Linear Filters**

Linear filters such as the Wiener filter in the wavelet domain yields peak outcomes when the signal corruption can be demonstrated as a Gaussian procedure and the accuracy criterion is the mean square error (MSE) [8]. It should be noted that creating a filter based on said statement often consequences in a filtered image that is more visually disagreeable than the original noisy signal despite the filtering operation successfully decreasing the MSE. In [9] a wavelet-domain spatially adaptive FIR Wiener filtering for image denoising is proposed where wiener filtering is performed only within each scale and intracule filtering is not allowed.

**II. Non-Linear Threshold Filtering**

The Non-linear Coefficient Thresholding based methods are among the most investigated domains through denoising through the use of (WT). The procedure showcases the property sparsity of the (WT) and the white noise maps in the signal domain. It is for that reason that when energy becomes more condensed into a lesser number of coefficients in the transform domain that the noise energy doesn't. Through this said mandatory principle that enabling the division between the signal and the noise.

The process where the small coefficients are detached whereas others remain untouched in what is commonly regarded to as Hard Thresholding. However, the method creates what is known as spurious blips or artifacts in the images as a result of unsuccessful attempts of removing temperately large noise coefficients. In order to get over the disadvantages of hard thresholding, (WT) employing soft thresholding was also presented in [4]. In such scheme, the coefficients above the threshold are shortened through complete value of the threshold itself. Just like soft thresholding, a variety of different techniques that apply thresholds are semi-soft thresholding and Garrote thresholding [5]. The majority of the wavelet shrinkage literature is based on methods that choose the optimum threshold which can be adaptive or non-adaptive to the image.

***a. Non-Adaptive thresholds***

VISU Shrink [7] is one of the universal non-adaptive thresholds, which rely solely on the amount of data points. It has asymptotic correspondence signifying the ideal performance as per the MSE when the number of pixels reaches infinity. VISU Shrink is known to yield overly smoothed images due to the choice of the threshold which can be unreasonably huge because of the requirement on the number of pixels within the image.

***b. Adaptive Thresholds***

A hybrid of the universal threshold is employed through SURE Shrink [7] and the SURE (Stein's Unbiased Risk Estimator) threshold and has better performance than VISU Shrink. Bayes Shrink [10] reduces the Bayes' Risk Estimator function with the assumption of Generalized Gaussian previously and so it yields data adaptive threshold. Bayes Shrink has been proven to work better than SURE Shrink in the majority of attempts. The wavelet coefficients are replaced with the weighted average of neighborhood coefficient to lower the Generalized Cross Validation (GCV) function in what would provide the maximum performance in Cross Validation [10].

Assuming that one can differentiate noise from the signal only according to the coefficient magnitudes is violated once the noise levels are higher than the signal scales. Under this high noise situation, the spatial formation of neighboring wavelet coefficients can perform a major role in noise-signal categorizations. Signals often shape significant features (e.g. straight lines, curves), while noisy coefficients often distribute at a random pace.

**III. Non-orthogonal Wavelet Transforms**

UDWT has been used for many applications including the decomposition of the signal to come up with better results and solutions. Since UDWT is shift invariant, it tends to avoid visual artifacts such as the pseudo-Gibbs phenomenon. Whereas the results have shown to have greatly improved, the usage of UDWT increases the overhead of computations and takes away from its feasibility in turn. Then through employing the Minimum Description Length principle that results showed that the Best Basis Function yielded the smallest code length required for the description of the given data. Then, thresholding was applied to denoise the data.

Additionally, the application of Multiwavelets is further explored to improve the performance, yet with increase in the computation complexity. The Multiwavelets are gathered through the application of over one mother function (scaling function) to the said dataset. Multiwavelets consist of some assets the likes of which are short support, symmetry, and one of the most crucial properties is the higher order of vanishing moments. This recipe of shift invariance and Multiwavelets is implemented in [10] where better results were given for the Lena image in context of MSE.

## IV. Wavelet Coefficient Model

The Wavelet Coefficient Model approach underlies the importance of showcasing the multiresolution assets of (WT). This method recognizes narrow correlation of signal at dissimilar purposes through detecting the signal across several resolutions. Output is found to be excelling in this method; however, it tends to be far more sophisticated in terms of computation and it is also found to be far more expensive. The modeling of the wavelet coefficients can either be deterministic or statistical.

### *a. Deterministic*

The Deterministic technique of modeling includes generating a tree structure of wavelet coefficients with each level in the tree presenting each scale of transformation and each of nodes representing the wavelet coefficients. This method was adopted in [5]. The best tree approximation shows a hierarchical clarification of wavelet decomposition. Wavelet coefficients of singularities have large wavelet coefficients that persist along the branches of tree. It is for that reason that if a wavelet coefficient has dominant presence at particular node then in the case of it being signal, its presence should be more pronounced at its parent nodes. If it is noisy coefficient, for instance spurious blip, then such consistent presence will be missing. The local maxima in scalespace was tracked by Lu et al. [11] through a tree structure. Donoho [12] proposed another wavelet coefficient tree for denoising.

### *b. Statistical Modeling of Wavelet Coefficients*

The Statistical Modelling of Wavelet Coefficients emphasizes on the properties of the (WT) including the multiscale correlation between the wavelet coefficients, the local correlation between neighborhood coefficients among other properties. The goal of this approach is to reach a perfect level of the exact modeling of image data through the use of (WT). There are plenty of reviews on the statistical properties of wavelet coefficients, two of them are found in [13]. The statistical properties are also exploited in the following techniques.

1. *Marginal Probabilistic Model:*

Numerous researchers have delved into the development of the massive local likelihood models for images in the wavelet domain. The marginal circulation of wavelet coefficients is highly kurtotic and often thus has a peak mark at zero and heavy tails. Two of the models that are most commonly used in modelling the wavelet coefficients distribution, namely: The Gaussian Mixture Model (GMM) [14] and the Generalized Gaussian Distribution (GGD) [15]. While the accuracy of GGD is higher, GMM provides a more simplified manner of use. Researchers have also recommended an approach where the wavelet coefficients which are presumed to be conditionally independent zero-mean Gaussian random variables, with the differences modeled as identically distributed, highly correlated random variables. An estimated Maximum A Posteriori (MAP) Probability rule is used for the approximation of marginal prior distribution of wavelet coefficient variances. The above-mentioned methods need a noise estimate, however, that could prove to be problematic to get in applications. A dual parameter Laplacian distribution was what Simoncelli and Adelson [17] employed for the wavelet coefficients of the image, which was is predicted from the noisy observations. Another method presented by Chang et al. [18] proposed using adaptive wavelet thresholding for image denoising through modeling the wavelet coefficients as a generalized Gaussian random variable, where parameters are locally calculated.

2. *Joint Probabilistic Model:*

The models of Hidden Markov Models (HMM) have proven to be efficient in capturing inter-scale dependencies, while the models of Random Markov Field [19] provide more efficiency in capturing the intrascale correlations. While difficulties of the local structures are not shown by Random Markov Gaussian densities, the Hidden Markov Models can be employed for capturing higher order statistics. The relationship between coefficients at matching scale but exist in a nearby neighborhood are modeled by Hidden Markov Chain Model while the correlation between coefficients across the chain is modeled by Hidden Markov Trees. Upon capturing the correlation, the HMM Expectation Maximization is employed for the approximation of the needed parameters and from which the denoised signal is predicted from noisy observation using a popular MAP estimator. One model that was described in [16] where each neighborhood of wavelet coefficients was described as a Gaussian Scale Mixture (GSM); which is one of the products of the Gaussian random vector and is a free hidden random scalar multiplier. In the study presented by Strela et al. [16], joint densities of cluster of wavelet coefficients as a GSM along with a maximum likelihood solution were presented for approximating the relevant wavelet coefficients from the noisy observations. In the work presented by Diwakar, Manoj, and Manoj Kumar [20], an approach for wavelet coefficients. One of the disadvantages of HMT is the computational load of the training stage. In order to overcome this computational obstacle, a basic HMT, named as uHMT was presented.

### 3.2.3 Data-Adaptive Transforms

In the recent years, a novel method was proposed under the name of ICA which has widely gained a lot of popularity and attention. The method has been employed successfully in [21] to denoise Non-Gaussian data. One of the excellent values of using ICA is its assumption of signal to be the Non-Gaussian that can aid in denoising images with Non-Gaussian along with the Gaussian distribution. Some of the disadvantages of ICA-based approaches in comparison to wavelet-based methods are the computational costs due to the usage of a sliding window and it needs a sample of noise free data or at least two image frames of the same scene. Some applications, however, may find it hard to gather the noise free training data.

## 4. CONCLUSION AND FUTURE WORK

The denoising algorithms performance is evaluated through the use of quantitative performance measures such as the Peak Signal-to-Noise Ratio (PSNR) and Signal-to-Noise Ratio (SNR) as well as in terms of visual quality of the images. A vast majority of the techniques that are currently in use assume the noise model to be Gaussian. However, the truth is that this assumption isn't always true due to the varied nature and sources of noise. An ideal denoising procedure requires a priori knowledge of the noise, whereas a practical procedure could possibly not have the necessary information about the variance of the noise or the noise model. It is for that reason that majority of the algorithms assume the identified difference of the noise and the noise model so that they could run a comparison on the performance with other algorithms. Gaussian Noise with different variance ideals is added in the natural images for the evaluation of the performance of the algorithm. Whilst not all of the researchers employ the usage of high value of variance in the evaluation of the performance of the algorithm when the noise is comparable to the signal strength. The use of FFT in filtering has been highly limited because of the limitations in providing scarce representation of data. The (WT) has proven to be the most suited for said performance due to its properties that include sparsity, multiresolution, and multiscale nature. Additionally, problems of computational difficulty should be taken in consideration. Thresholding techniques used with the DWT are known to be the simplest especially when it comes to implementation. Non-orthogonal wavelets such as UDWT and Multiwavelets have a tendency of improving the level of the performance; however, that comes at the price of a large overhead in their computation. HMM based methods seem to be promising; yet they are quite complicated. Upon using using (WT), Nason [22] highlighted that one of the main problems is the choice of primary resolution and analyzing wavelet. For they have a major influence on the success of the shrinkage procedure. When the algorithms are compared, researchers emphasize the importance on not omitting the details of said comparison. Plenty of works did not specify the wavelet used neither the level of decomposition of the (WT) was mentioned.

Future works and research are prophesized to focus on the development of robust statistical models of non-orthogonal wavelet coefficients as the intra and inter scale correlations. Those models could be used with efficiency for image denoising and compression.

## References

[1] Mahdianpari, Masoud, Bahram Salehi, and Fariba Mohammadimanesh. "The effect of PolSAR image de-speckling on wetland classification: introducing a new adaptive method." Canadian Journal of Remote Sensing 43.5, pp.485-503, 2017.
[2] S. G. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," IEEE Trans. Inform. Theory, vol. 38, pp. 617-643, 1992.
[3] D. L. Donoho, "De-noising by soft-thresholding.", IEEE Trans. Information Theory, vol.41, no.3, pp.613-627, 1995.
[4] Hend A., Omar G., Shawkat K. "Volumetric Medical Images Lossy Compression using Stationary Wavelet Transform and Linde-Buzo-Gray Vector Quantization.", Research Journal of Applied Sciences, Engineering and Technology, vol.14, no.9, pp. 352-360, 2017.
[5] Omar G., Mahmood A., Hend A., Shawkat K. "Hybrid Compression based Stationary Wavelet Transforms.", International Journal & Magazine of Engineering, Technology, Management and Research, vol.3, no.11, pp. 524-527, 2016.
[6] Shivaleela, Manchala, L. N. Anitha, and Verilog Engineer. "A Novel Method for Noise Removal in Images Using Bit Planes Through Standard Mean Filter.", International Journal of Advanced Research in Electronics and Communication Engineering, vol.1, no.1, pp. 30-34, 2018.
[7] Zhou, Pan, et al. "Dictionary learning with structured noise." Neurocomputing 273, pp.414-423, 2018.
[8] Omar G. "Lossy Compression Using Stationary Wavelet Transform and Vector Quantization.", Master of Thesis, Alexandria University, 2016.
[9] H. Zhang, Aria Nosratinia, and R. O. Wells, Jr., "Image denoising via wavelet-domain spatially adaptive FIR Wiener filtering", in IEEE Proc. Int. Conf. Acoust., Speech, Signal Processing, Istanbul, Turkey, June 2000.
[10] Goyal, Bhawna, et al. "Two-dimensional gray scale image denoising via morphological operations in NSST domain & bitonic filtering." Future Generation Computer Systems, 2018.
[11] J. Lu, J. B.Weaver, D.M. Healy, and Y. Xu, "Noise reduction with multiscale edge representation and perceptual criteria," in Proc. IEEE-SP Int. Symp. Time- Frequency and Time-Scale Analysis, Victoria, BC, pp. 555-558, 1992.
[12] D. L. Donoho, "CART and best-ortho-basis: A connection," Ann. Statist., pp. 1870-1911, 1997.
[13]. Chris W. "Beyond ICA.", School of Informatics, University of Edinburgh, pp.1-25, 2018.
[14]. Goyal, Bhawna, et al. "Two-dimensional gray scale image denoising via morphological operations in NSST domain & bitonic filtering." Future Generation Computer Systems, 2018.
[15]. Deledalle, Charles-Alban, Shibin Parameswaran, and Truong Q. Nguyen. "Image restoration with generalized Gaussian mixture model patch priors." arXiv preprint arXiv:1802.01458, 2018.
[16] V. Strela, J. Portilla, and E. P. Simoncelli, "Image denoising via a local Gaussian scale mixture model in the wavelet domain," in Proc. SPIE 45th Annu. Meeting, San Diego, CA, 2000.

[17] E. P. Simoncelli and E. Adelson, "Noise removal via Bayesian wavelet coring," in Proc. IEEE International Conference on Image Processing, Lausanne, Switzerland, September, pp. 279-382, 1996.
[18] S. G. Chang, B. Yu, and M. Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising." IEEE Trans. Image Processing, vol. 9, pp. 1522-1531, 2000.
[19] Green, Michael, et al. "A Neural Regression Framework for Low-Dose Coronary CT Angiography (CCTA) Denoising." International Workshop on Patch-based Techniques in Medical Imaging. Springer, Cham, pp. 102-110, 2017.
[20] Diwakar, Manoj, and Manoj Kumar. "A review on CT image noise and its denoising." Biomedical Signal Processing and Control 42, pp.73-88, 2018.
[21] Kamkarhaghighi, Mehran, and Masoud Makrehchi. "Content Tree Word Embedding for document representation." Expert Systems with Applications 90, pp. 241-249, 2018.
[22] Nason, G.P., "Fast cross-validatory choice of wavelet smoothness, primary resolution and threshold in wavelet shrinkage using the Kovac- Silverman algorithm", Technical report, Department of Mathematics, University of Bristol, United Kingdom, 1988.