Medical Image Denoising Using Two-Stage Iterative Down-Up CNN and SURF Features

M.R. Jospine Shiney¹, PG Student, Hepsiba D², L. D. Vijay Anand³, Assistant Professor², ³

Department of biomedical engineering, Karunya Institute of Technology and Sciences, Coimbatore, Tamil Nadu, India.

Email: jospineshiney@karunya.edu.in¹, hepsiba@karunya.edu², vijayanand@karunya.edu³

ABSTRACT. This paper proposes a Two-Stage Deep Iterative Down-Up Convolutional Neural Network for denoising the medical images, which more often increase and decrease the feature map resolution. The function, Speeded Up Robust Features (SURF) is employed in the proposed system to handle the problem of gradient vanishing. The SURF is a patented local feature detector and descriptor. Experiment is conducted using several different medical noise images like Computed Tomography (CT) and Ultrasound Image and found that the proposed system outperforms the current state-of-the-art consistently in image denoising methodologies.

Keywords. Image Denoising, Image Decomposition, Medical Noise Image, SURF Features, Iterative Down-Up Convolutional Neural Network

I. INTRODUCTION

The medical noises are the great barriers in the measurement procedures that causes quality degradation and uncertainty in the diagnosis which makes those images unusable. There are several existing approaches available for medical image denoising over the past decades. The noises are of different for different medical imaging instruments and are classified as Fixed pattern noise and Independent random noise.

Speckle is a multiplicative noise which means the granular pattern that occurs on B-scans by mottling. The phased array ultrasound scanners are of coherent in nature. Hence, the indeterminable random tissue in homogeneities causes mottled images of B-scan images of kidney and liver, as it has too small underlying structures that can be resolved by ultrasound scan. The quality of the B-scan images could be degraded by speckle which in turn minimizes the human perception to determine the fine details in those scan images during diagnosis. Also, it results in loss of efficiency in edge detection process [1]. The ultrasound imaging when invented, it goes on increasing in its usage in real-time which is considered as a most crucial technique in medical diagnosis. For an ultrasonic imaging to be successful in its diagnosis, image quality is the important feature. But this image is often affected by the speckle noise, which is a form of acoustic noise that degrades the ultrasound image quality more. The speckle noise can also be seen in the laser and microwave radar imaging as interference effect that occurs due to the scattering of ultrasonic beam arises from the microscopic tissue inhomogeneity. The granular pattern obtained will not be related to the actual tissue micro
structure. But the speckle covers the availability of lesions in the images and minimizes the human perception towards those minute details. Thus, it is important to suppress the speckle noise in medical ultrasound images to enhance the quality of those images for successful diagnosis of fine details in it [2]. The CT is a technique of radiographic examination. It takes in the large series of 2D image of an object’s cross-section and generates a 3D image of the same object. The CT is used in many of the clinical situations connected to conventional radiography. The major disadvantage of using conventional radiography techniques that depicts 3D objects as 2D images are, the overlaid on the image. This issue is solved in CT due to the scanning of thin slices of body using a narrow beam of x-rays that are rotated over the stationary patient’s body [3].

There are several automotive, gaming and entertainment applications that contain 3D vision. Here, the estimation of distance between the sensor and the object is very important. This technique is called as range or depth sensing and can be done with structured light, LiDAR, depth from stereo and Time of Flight (ToF) active sensing. The ToF is now in trending due to the fact that it can deliver the scene geometry details in 2D map format ie., range, depth. This can then be interpreted as 2D gray scale image [4]. In digital cameras and camcorders, more sophisticated algorithms are embedded for the sake of image sharpening and noise reduction processes. The known fact is that, the contrast enhancement increases the noise occurred due to sharpening, which is the major disadvantage of several existing contrast enhancement algorithms. The impact of this drawback is deceptive towards the classical linear unsharp masking process, in which a part of image undergone high pass filtering is added to the input data. If this high pass filter is replaced by the non-linear filters, then there could be a specific improvement, as it combines the edge enhancement and noise reduction effectively, resulting in limited noise amplification while sharpening an image [5]. Breast cancer is now common among women which causes to death. The National Cancer Institute reported that among one eighth of the women of United States are affected by this disease. The cause of breast cancer to be developed in someone is still unknown and hence, the primary prevention appears to be impossible. But the detection of this disease in its early stage can be a key factor that improves the prognosis of breast cancer. The most popular procedure of screening and diagnosing breast cancer is, X-ray mammography. The breast cancer when detected earlier can reduce the mortality rate upto 25%. But around 10-30% of breast lesions cannot be interpreted or can be missed in screening process [6].
in transmission and acquisition process, the medical images are corrupted by noise, which reduces the fusion effect. The medical images with noise lead to mistakes in characterizing the images and hence, the noise in a medical image is a big challenge in many traditional methods of image fusion [9].

The denoising an image is a vital process, since the denoised image can be used for several useful applications. Moreover, image denoising can be used in all the fields where image processing is assessed, as it is a simplest possible inverse problem [10]. The signal denoising is most important in the field of measurement and instrumentation. This is because; it reduces the uncertainties in measuring process and enhances the accuracy and credibility of several procedures. Methods like adaptive filters, statistical estimators, transform domain methods etc. had been explored to address this problem. In recent, the sparse representation theory is a growing technique, which suggests that the clean parts of a signal with noise can have a good sparse representation that makes use of a pre-defined dictionary, but the dictional can’t sparsely represent the noisy parts [11]. The sparse representation can be a better tool for image fusion, image denoising, compressive sensing, electroencephalogram inverse problem, voxel selection, magnetic resonance spectroscopy, etc. The traditional sparse representation systems assume that, the non-zero coefficients could be appeared in random. But the sparse coefficients show intrinsic structures often in cluster form. The standard sparse representation won’t consider intrinsic sparse signal. Hence, integrating the intrinsic structures in sparse representation is a technique for enhancing the sparse representation performance [12].

Based on the transform and signal property, the sparsity of the representation varies. Good sparsity can be achieved for multiresolution transforms for spatially localized information like edge and singularity, since they are more available in natural images and displays a substantial part of the data into it. Hence, this transform can be applied significantly in image denoising methodologies [13]. For denoising the Magnetic Resonance Spectroscopic Imaging (MRSI) signals, there are various methods available. This signal can be denoised by consecutive projection against various areas characterized by a set of linear time-frequency transform. Explicit parametric models can also be employed in MRSI denoising either before or during the metabolite quantitation. For correct models, this method is much effective and for incorrect models, bias problem will occur [14]. Remote sensing imagery is a significant technology that can be used in several different types of applications like earth climate, military and agriculture. But practically, different types of stripe noise would be present in the remote sensing imagery together with the cross-track and push-broom imaging. This is because of the variations in calibration error, detector response, etc. The remote sensing images with stripe noise degrade the quality of the images visually and risk the suitability of those images for further processing [15].

A technique called Industrial tomography, determines the distribution of the materials in an imaging region with respect to the set of multi-angle measurements. There are 12 electrodes being embedded onto the outer wall of the tube in Electrical Capacitance Tomography (ECT). The capacitance between the electrode pairs is calculated which is then used for reconstruction of images. The inverse problem is needed to be solved in this purpose [16]. A technique called Single Image Super-Resolution (SISR) is used to generate an image of high-resolution from low-resolution. This technique can be used more commonly in the field of computer vision applications like medical, surveillance and security imaging etc. where there is a demand for more fine details from an image is required. Traditional SISR method includes interpolation
like bicubic interpolation and Lanczos resampling, methods that use internal patch recurrence or statistical image priors [17]. In medical and some other industries, XRAY-CT is used widely. The increased use of medical computed tomography leads to excessive exposure of the patients to radiation dosage. Based on the ALARA (As Low As Reasonably Achievable) principle, extensive research is being undergone in reducing the dosage of CT [18].

For scanning a patient with X-ray CT scan in routine, they are supposed to be developing the risk of cancer due to the exposure to the radiation more often. Hence, it is important to minimize the routine scanning in patients. But the low dose of X-ray CT causes severe artifacts because of the less number of photons, beam hardening etc. that leads to low reliability of diagnosis. Therefore, the interest in research of reconstructing a high-quality data from low dosage of X-ray CT is increasing in the community of CT. Conventional denoising methodologies are high in computation expenses and the image denoising methods cannot eliminate the CT-specific noise more readily. To overcome these issues, a deep learning algorithm based low dose X-ray CT is introduced [19]. CT, Magnetic Resonance Imaging (MRI), Medical imaging including X-rays, ultrasound etc. are more vulnerable to noise. The cause for this noise varies with respect to image acquisition methods and reduced exposure of patients to radiation, i.e., if the time taken to expose the patient to radiation is decreased, then the noise will get increased. Hence, denoising is vital in this situation for appropriate image analysis by machines and by the humans [20]. The contributions in the proposed medical image denoising method based on two-stage iterative down-up CNN and SURF features are briefed as follows:

- The characteristics of noise and image are considered from the perception of image decomposition. The proposed system restores the desired image by leveraging the noise map which makes the proposed method more robust towards various noise levels.

- The propose two-stage iterative down-up CNN model can process several different types of medical noises faster and with better performance. Experiment is conducted using several different medical noise images like CT and Ultrasound Image and found that the proposed system outperforms the current state-of-the-art consistently in image denoising methodologies by means of effectiveness and efficiency.

2. RELATED WORKS

In medical B-scan imaging, a speckle suppression based on adaptive smoothing is introduced. In this method, the local kernels with appropriate shape and size are filtered. A filtered kernel can be obtained from each of the pixels that gets fitted onto local homogeneous area with processed pixels. This can be done by the region growing technique based on local statistics. This model is then examined using phantom and images and is found that the speckle can be reduced effectively by this filter, yet the resolvable information would be preserved [1]. A space-invariant filter cannot perform a single operation on every pixel of an image. Also, in nonlinear edge-preserving filters, same problem occurs. Hence, a space-varying algorithm that makes use of local image content can be used to solve this problem [2]. Several noise reduction filter to reduce the noise in CT images is examined. The combination of Gaussian and Prewitt operators and on anisotropic diffusion-based filters is used in [3].
The modeling and removal of fixed pattern in photonic mixture devices that makes use of the principle called Time-of-flight for calculating the range and depth is used in [4]. The filtering of input image based on 2 types of directional smoothers called Type 1 and Type 2 PWL Directional Smoothing are used. Sharpening module is used to process the results of the filter. Based on the value of Gaussian noise, the amount of required smoothing and sharpening would be chosen [5]. A dyadic wavelet processing-based image denoising and image enhancement algorithm is demonstrated which makes use of the estimation of local iterative noise variance [6].

A wavelet transform-based hyper-analytic wavelet transform combined with other filters used in discrete wavelet transform is introduced and this method is a simple, yet faster image denoising model [7]. A variation-based model that makes use of sparse prior to remove ring artifacts in the CT slice. With the help of Alternating Direction Method of Multipliers (ADMM), the model is divided into many simple sub-problems that can be solved easily [8]. A variation-based model that can perform multi-modal medical image fusion and medical image denoising is presented. This system uses multiscale alternating sequential filter for image fusion and Adaptive fractional order total variation (AFOTV) constraint to denoise the medical images [9].

The zero-mean white and homogeneous Gaussian additive noise is removed from the input image by using K-Singular Value Decomposition (K-SVD) algorithm based on the redundant representation and sparse representation over trained dictionary is given in [10]. A Random Refined Orthogonal Matching Pursuit (RROMP) algorithm, which is an algorithm of sparse recovery, is used for denoising the signal by generating many sparse representations using False Discovery Rate (FDR) control and multi-selection strategy [11]. A Dictionary Learning method based on Group Sparsity and Graph Regularization (DL-GSGR) is introduced. It takes a 3D medical image for denoising it using group sparse representation [12].

An image denoising method that uses enhanced sparse representation in the transform domain is described. In this method, same types of 2D fragments of an image are grouped into 3D arrays. This 3D group is processed by Collaborative Wiener Filtering [13]. The MRSI images with high Signal to Noise Ratio (SNR) are denoised using Casorati and Hankel form of matrix and low-rank approximation through Singular Value Decomposition (SVD). This method can be applicable in denoising the MRSI images, spatial-temporal images and spatial-spectral also [14]. A Low Rank Based Single Image Decomposition (LRSID) model is employed to naturally convert the image de-striping problem to image decomposition problem and separating the original image from the stripes [15].

A low-rank image decomposition method is demonstrated to reconstruct the images of ECT robustly and to recover the corrupted and missing pixels in it. To get the exact error matrix and low rank matrix, convex optimization is carried out [16]. A Single Image Super Resolution model based on very deep convolutional neural network of VGG net is established to propose an effective and simple training process for Image Net classification [17]. The deconvolution network, auto-encoder and short-cut connections are combined into a Residual Encoder-Decoder Convolutional Neural Network (RED-CNN) for low dose CT imaging based on deep learning is introduced in [18].

A CNN model over the coefficient of wavelet transform of low-dose CT image is described in [19] for the extraction of directional component of artifacts and the exploitation of intra band and inter band correlations to minimize the CT image noise. The auto-encoders based on
denoising with convolutional layers is demonstrated for medical image denoising efficiently by boosting the sample size by combining the heterogeneous images to enhance the performance [20].

3 SYSTEM METHODOLOGY

The proposed model uses a Deep Two-Stage Iterative Down-Up CNN for medical image denoising, that more often causes variations in the of feature map resolution and also the Speeded Up Robust Features (SURF) which extracts, detects and describes the useful features in the medical images in a dataset required for successful image denoising.

(i). Deep Learning Based Image Denoising

The deep learning approach when integrated into medical image denoising improves the performance of the denoising algorithm. Several image denoising algorithms are available and some of them are described as follows. A simple CNN architecture with 5 denoising layers cannot achieve significant performance. An image denoising model based on auto encoder is also observed to have low performance when compared to Block Matching and 3D Filtering (BM3D). A multilayer perceptron is introduced to learn mapping from noised image patches to denoised image, attains a performance which is similar to that of the BM3D. The performance of the Denoising Convolutional Neural Network (DCNN) is improved when compared to conventional models using convolution, batch normalization and Rectified Linear Unit (ReLU) as basic structures and trained deep networks by using global residual learning. A CNN model with 7 layers is introduced that considers the trade-off among the accuracy and the computational cost. This is successful in enhancing the receptive field using dilated convolution. The Memory Network (MemNet) provides a larger receptive field while sustaining less parameter with the help of a Recursive CNN and dense skip connections. A Residual Dense Network (RDN) which makes use of residual learning and dense connection as basic structure, increases the reuse of the features and attains considerable performance enhancement in image denoising using Gaussian noise. All these methods have its own limitations and to overcome those, a deep network based on Up-Down scaling is proposed in this paper.

(ii). Deep Networks Using Up-Down Scaling

The depth and the computational complexity of a network can be maintained when the receptive field gets increasing, a dilated convolution is used which is affect by gridding artifacts, but it can sub-sample the features sparsely. There should be a better trade-off between the computational cost and the receptive field, hence down-scaling and up-scaling of feature maps should be improved. The resolution of feature maps can be reduced to half by using max-pooling, simultaneously the number of feature maps can be increased by a factor of 2 to minimize the data loss. The Up-sampling with 2x2 convolution is used in up-scaling the low-resolution features. A method called Wavelet Transform (WT) that consists of convolution and sub-sampling can be used for down-scaling and up-scaling of the feature maps.
Fig. 1: Architecture of Deep Iterative Down-Up Network (DIDN) Model

The deep down-up network comprises of feature extraction, Down-Up Block (DUB), reconstruction and enhancement. The feature extraction is carried out using SURF. Consider the input image size is \( X \times Y \), then the DIDN extracts \( N \) features at first using 3\( \times \)3 convolution, and feature extraction is performed by the size \( \frac{X}{2} \times \frac{Y}{2} \times 2N \) via convolution layer of 2 strides. The feature that are extracted from the input image are then given to DUB, in which two up and down scaling processes are performed. For down-scaling, a 3\( \times \)3 convolution layer with 2 strides is used and for up-scaling, a sub-pixel layer is used. The reconstruction block comprises of 9 convolution layers and a Parametric Rectified Linear Unit (PReLU). The enhancement block uses 1\( \times \)1 convolution to reduce the number of feature maps in the output of reconstruction block, then an up-scaling process is done in the sub-pixel layer to produce a denoised image as an output.

(iii). Two-Stage Convolutional Neural Network

The propose deep iterative down-up CNN iteratively down and up scales the feature maps with the help of a convolutional layer and a deconvolutional layer by 2 strides.

a. Dataset

In this paper, 2 different types of dataset are used. The different medical images collected for developing a medical image dataset for image denoising process includes, the images from devices like Computed Tomography (CT) and Ultrasound Imaging. The CT images dataset is obtained from National Biomedical Imaging Archive (NBIA) and the Ultrasound images dataset is obtained from the teaching files of Gelderse Vallei Hospital in Ede, Netherlands. This
dataset consists of a larger number general cases of ultrasound imaging collected over years by the ultrasound technicians and the radiologists form that hospital.

b. Image Decomposition
The noises in the medical images mostly consist of random and structural noises. Hence image degradation can be done by using the following expression:

\[ Y = X + N \]

Where
- \( X \) - Ground truth image,
- \( Y \) - Input noise image,
- \( N \) - Noise.

The image decomposition process estimates simultaneously the values of \( X \) and \( N \) from the value of \( Y \).

c. SURF Feature Extraction
The SURF is used to extract the local robust features based on the Hessian matrix for detecting and based on Hessian distribution for describing the features. The SURF uses Box Filter to approximate the Laplacian of Gaussian (LoG). This approximation provides an advantage of easier calculation of convolution with box filter using the integral images parallelly done different scales. In addition to this, the SURF depends on the Hessian matrix determinant for location and scales.

HESSIAN DETECTOR
The Hessian Detector detects the important features automatically regardless of variations in view-points. This can be generally described as the process of attaining the scale invariance for finding verifying the scale features at different scales or scale space that can be attained by splitting the scale space into various levels and octaves. The scale space is implemented in the form of image pyramids as given in the Fig. 3. The levels of the pyramid are found by the Gaussian kernels sub-sampling and smoothing. The SURF makes use of a Hessian detector for subtracting those layers of the pyramid. The scale space can be evaluated by up-scaling the filter sizes of the integral image combined to fast Hessian matrix method. The processing time of the SURF filters are size invariant. Hence the simultaneous processing is possible in SURF and need for image sub-sampling can be negotiated thereby improving the performance. A Hessian-based blob detector is used by SURF for finding the important features. The Hessian matrix determinant provides the extent of response and it is the expression of local changes across the region.
Fig 2: $L_{xx}$ and $L_{yy}$ Discrete Gaussian and its approximations $D_{xx}$ and $D_{yy}$.

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix}$$

where,

$$L_{xx}(x, \sigma) = I(x) \ast \frac{\partial^2 y}{\partial x^2} g(\sigma)$$

$$L_{xy}(x, \sigma) = I(x) \ast \frac{\partial^2 y}{\partial x \partial y} g(\sigma)$$

where,$L_{xx}(x, \partial)$ - convolution of an image through 2nd order Gaussian derivative, $g(\partial)$,

$I(x)$ - integral image where each point $x = (x, y)^T$ stores the sum of all pixels within a rectangular region.

The major need for SURF detection is the non-maximal suppression of determinants in Hessian matrix. The cost of convolution is high and it uses approximations and can be speeded-up by using integral images and approximated kernel.

$$I(x) = \sum_{i=0}^{sx} \sum_{j=0}^{sy} I(x, y)$$
The SURF approximates the kernels using the rectangular box, box filter. Hence, the calculation of approximated convolution can be effective for arbitrary sized kernels that uses integral image.

\[
D_{et}(H_{app}) = D_{xy}D_{yy} - (wD_{xy})^2
\]

The approximate and discrete kernels are termed as \(D_{yy}\) for \(L_{yy}\) and \(D_{xy}\) for \(D_{xy}\). The term, \(w\) is scale-sensitive but remains as constant at 0.9.

**Fig 3: Integral image usage up-scales the filter at constant cost**

**DESCRIPTOR**

A descriptor must provide robust and unique descriptions of features. It is generated by the region around the interest points. The SURF descriptor is uses Haar wavelet response which is efficiently calculated by using the integral images. To achieve a rotational invariance, unique orientation for a feature is needed to be determined. In this case, before calculating the descriptor, the interest point surround region should be rotated to its direction.

**Fig 4: Descriptor Components**

The SURF descriptor defines the region of interest of size 20s. The region of interest is split into 4x4 sub-regions as defined by the wavelet response values in the directions of \(x\) and \(y\). The wavelet response in the direction, \(x\) and \(y\) is denoted respectively as \(d_x\) and \(d_y\). For each of the sub-region, a vector, \(v\) is computed on the basis of 5x5 samples.

\[
v = \{ \sum d_x, \sum |d_x|, \sum d_y, \sum |d_y|, \}
\]
In Fig. 4, the selected square limits one of the 16 sub-regions and the dots characterizes the sample points. The descriptor of the features are 16-vectors for concatenated sub-regions. It is then undergone normalization to get the invariance to contrast variation and therefore represented as the linear scaling of the descriptor.

d. Two-Stage Iterative CNN (TSIC)

In this paper, a model of Two-Stage Deep Iterative Down-Up CNN is proposed for denoising the medical images and makes them suitable for diagnosis. This network consists of 3 layers namely, convolution, batch normalization and Rectified Linear Unit (RELU) layers. The convolutional layer extracts the features using SURF. The batch normalization layer prevents the gradient vanishing divergence problem. The RELU layer is used to track the sparsity and its nonlinearity. The algorithms for training and testing process of medical image denoising using two-stage iterative Down-Up CNN model are given in Algorithm-1 and Algorithm-2 respectively.

Fig. 5: Proposed TSIC Model
Algorithm 1: Training

Initialize:
Total number of images in a folder as total_images
Directory of keyframe as image_folder
Read images as filenames
for n = 1: total_images
    Label the images,
    full_name = image_folder, filenames
    Denoise the images
    Add Gaussian noise
end
l=0;
Define a starting folder
    Start_path = folder_directory
Folder Confirmation/changing,
    topFolder = start
if topFolder == 0
    return
end
Get subfolders list,
    SubFolder = genpath (topFolder)
Parsing cell array,
    remaining = SubFolder
    FolderName = {}
while true
    [singleSubFolder, remaining] = strtok (remaining, ';')
    if isempty (singleSubFolder)
        break
    end
    FolderName = [FolderName singleSubFolder]
end
Get list of all files in the folder,
    I2=mydata{k}
SURF Feature Extraction
    points = detectSURFFeatures (rgb2gray (I2))
if (l==1)
    Noised
end
if (l==2)
    Denoised
end
Load Model,
    net = resnet50()
Split dataset into Training set and Test set,
Convert matrix to grey image
Resize the image

Algorithm 2: Testing

Initialize:
trainingLabels = trainingSet.Labels
net = denoisingNetwork ('DnCNN')
Pass CNN image features to trained classifier
Get the known labels,
    testLabels = testSet.Labels
Input:
I3 = Original Image
Resize original image,
I3 = imresize (I3, [200 200]);
Convert to gray
Histogram of grey image,
figure,imhist (gray);
SURF Feature Extraction,
points = detectSURFFeatures (rgb2gray (I3))
denoisedImage = denoiseImage (gray,net)
Compute PSNR
Compute Accuracy, Sensitivity, Specificity, Precision, F_Score
end

4. PERFORMANCE MEASURES
The various performance metrics used in the measurement of the performances of the proposed methodology over different dataset are given as follows:

(i). Accuracy
The accuracy of a propose model can be evaluated by calculating the average of the True values in the result obtained from the given dataset of medical images and is referred to as the ratio of sum of the True Positives and Negatives to the total of the True Positives and Negatives and the False Positives and Negatives.

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} * 100
\]

where,
TP – True positive
TN – True Negative
FP – False Positive
FN – False Negative

(ii). Sensitivity
True Positive Rate (TPR) is referred to as the ratio of True Positive to sum of True Positive and False Positive.

\[
\text{TPR} = \frac{TP}{TP + FN}
\]

(iii). Specificity
True Negative Rate (TNR) is referred to as the ratio of True Negatives to sum of the True Negatives and False Positives.

\[
\text{TNR} = \frac{TN}{TN + FP}
\]

(iv). Miss Rate
False Negative Rate (FNR) is the ratio of False Negatives to sum of the False Negatives and True Positives.
(v). **Fall Out**

False Positive Rate (FPR) is the ratio of False Positives to sum of the False Positives and True Negatives.

\[
FPR = \frac{FP}{FP + TN}
\]

(vi). **Recall**

The Recall is referred to as the ratio of True positives to sum of the True Positives and False Negatives obtained from the dataset of medical images.

\[
Recall = \frac{TP}{TP + FN}
\]

(vii). **Precision**

The performance metric Precision is referred to as the ratio of True positives to sum of the True and False Positives obtained from the dataset of medical images.

\[
Precision = \frac{TP}{TP + FP}
\]

(viii). **F1-Score**

The F-measure, \(F_\beta\) is computed by using the Precision and Recall metrics. The value of \(F_\beta\) provides the metric F1-score. F1-score is referred to as the Harmonic mean of the values obtained from the Precision and the Recall.

\[
F_\beta = \frac{(1 + \beta^2)(Precision \times Recall)}{\beta^2 \times (Precision + Recall)}
\]

Substitute \(\beta = 1\), after simplification,

\[
F_\beta = \frac{2 \times (Precision \times Recall)}{1 \times Precision + Recall}
\]
5 Results and Discussion

Fig. 6: Original Input Image

Fig. 7: Histogram of Original Image
Fig. 8: Extracted SURF Features

Fig. 9: Original Noisy Image (left) and Denoised Image (right)

Fig. 10: Accuracy Plot
6. CONCLUSION

In this paper, a Two-Stage Deep Iterative Down-Up Convolutional Neural Network is proposed for denoising the medical images, which more often increase and decrease the feature map resolution. In contrast to the existing approaches of medical image denoising, the proposed approach utilizes the characteristics of both image and noise components simultaneously for processing. The Speeded Up Robust Features (SURF) is used in the proposed system which acts as a feature detector and a feature descriptor. Extensive experiments are conducted using various medical noise image datasets like Computed Tomography (CT) and Ultrasound Image and found that the proposed model consistently exceeds the performance of the current state-of-the-art in medical image denoising efficiently and effectively. The major advantages of the proposed model are the speed and latency, security and cost savings. The speed and latency mean the traffic loads of a larger enterprise in computing the medical data are reduced thereby improving the performance if the applications and the services. The security means the problems related to the local compliances, privacy regulations and the data sovereignty are resolved. Cost savings means the computation not only eliminates the need for cloud but also optimizes the data flow which increase the computational costs.

REFERENCES

[1] M. Karaman, M. A. Kutay, and G. Bozdagi, “An adaptive speckle suppression filter for medical ultrasonic imaging,” IEEE Trans. Med. Imag., vol. 14, no. 2, pp. 283–292, 1995.
[2] T. Loupas, W. McDicken, and P. L. Allan, “An adaptive weighted median filter for speckle suppression in medical ultrasonic images,” IEEE Trans. Circuits and Syst., vol. 36, no. 1, pp. 129–135, 1989.
[3] F. Attivissimo, G. Cavone, A. M. L. Lanzolla, and M. Spadavecchia, “A technique to improve the image quality in computer tomography,” IEEE Trans. Instrum. Meas., vol. 59, no. 5, pp. 1251–1257, 2010.
[4] M. Georgiev, R. Bregović, and A. Gotchev, “Fixed-pattern noise modelling and removal in time-of-flight sensing,” IEEE Trans. Instrum. Meas., vol. 65, no. 4, pp. 808–820, 2016.
[5] F. Russo, “An image-enhancement system based on noise estimation,” IEEE Trans. Instrum. Meas., vol. 56, no. 4, pp. 1435–1442, 2007.
[6] A. Mencattini, M. Salmeri, R. Lojacono, M. Frigerio, and F. Caselli, “Mammographic images enhancement and denoising for breast cancer detection using dyadic wavelet processing,” IEEE Trans. Instrum. Meas., vol. 57, no. 7, pp. 1422–1430, 2008.
[7] I. Firoiu, C. Nafornta, J.-M. Boucher, and A. Isar, “Image denoising using a new implementation of the hyperanalytic wavelet transform,” IEEE Trans. Instrum. Meas., vol. 58, no. 8, pp. 2410–2416, 2009.
[8] L. Yan, T. Wu, S. Zhong, and Q. Zhang, “A variation-based ring artifact correction method with sparse constraint for flat-detector CT,” Phys. Med. Biol., vol. 61, no. 3, p. 1278, 2016.
[9] W. Zhao and H. Lu, “Medical image fusion and denoising with alternating sequential filter and adaptive fractional order total variation,” IEEE Trans. Instrum. Meas., vol. 66, no. 9, pp. 2283–2294, 2017.
[10] M. Elad and M. Aharon, “Image denoising via sparse and redundant representations over learned dictionaries,” *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, 2006.

[11] S. Li and L. Fang, “Signal denoising with random refined orthogonal matching pursuit,” *IEEE Trans. Instrum. Meas.*, vol. 61, no. 1, pp. 26–34, 2012.

[12] S. Li, H. Yin, and L. Fang, “Group-sparse representation with dictionary learning for medical image denoising and fusion,” *IEEE Trans. Biomed. Eng.*, vol. 59, no. 12, pp. 3450–3459, 2012.

[13] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image denoising by sparse 3-d transform-domain collaborative filtering,” *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, 2007.

[14] H. M. Nguyen, X. Peng, M. N. Do, and Z.-P. Liang, “Denoising MR spectroscopic imaging data with low-rank approximations,” *IEEE Trans. Biomed. Eng.*, vol. 60, no. 1, pp. 78–89, 2013.

[15] Y. Chang, L. Yan, T. Wu, and S. Zhong, “Remote sensing image stripe noise removal: from image decomposition perspective,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7018–7031, 2016.

[16] J. Ye, H. Wang, and W. Yang, “Image recovery for electrical capacitance tomography based on low-rank decomposition,” *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1751–1759, 2017.

[17] J. Kim, J. Kwon Lee, and K. Mu Lee, “Accurate image super-resolution using very deep convolutional networks,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 1646–1654.

[18] H. Chen, Y. Zhang, M. K. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, and G. Wang, “Low-dose CT with a residual encoder-decoder convolutional neural network,” *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2524–2535, 2017.

[19] E. Kang, J. Min, and J. C. Ye, “A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction,” *Med. Phys.*, vol. 44, no. 10, 2017.

[20] L. Gondara, “Medical image denoising using convolutional denoising autoencoders,” in *Proc. Int. Conf. Data Mining Workshops*, 2016, pp. 241–246.