Noise-residue learning convolutional network model for magnetic resonance image enhancement

Ram Singh¹ and Lakhwinder Kaur²

Department of Computer Science & Engineering, Punjabi University Patiala - 147002 (Punjab), INDIA
E-mail: bhankharz@gmail.com, mahal2k@yahoo.com

Abstract. Magnetic Resonance Image (MRI) is an important medical image acquisition technique used to acquire high contrast images of human body anatomical structures and soft tissue organs. MRI system does not use any harmful radioactive ionized material like x-rays and computerized tomography (CT) imaging techniques. High-resolution MRI is desirable in many clinical applications such as tumor segmentation, image registration, edges & boundary detection, and image classification. During MRI acquisition, many practical constraints limit the MRI quality by introducing random Gaussian noise and some other artifacts by the thermal energy of the patient body, random scanner voltage fluctuations, body motion artifacts, electronics circuits impulse noise, etc. High-resolution MRI can be acquired by increasing scan time, but considering patient comfort, it is not preferred in practice. Hence, post-acquisition image processing techniques are used to filter noise contents and enhance the MRI quality to make it fit for further image analysis tasks. The main motive of MRI enhancement is to reconstruct a high-quality MRI while improving and retaining its important features. The new deep learning image denoising and artifacts removal methods have shown tremendous potential for high-quality image reconstruction from noise degraded MRI while preserving useful image information. This paper presents a noise-residue learning convolution neural network (CNN) model to denoise and enhance the quality of noise-corrupted low-resolution MR images. The proposed technique shows better performance in comparison with other conventional MRI enhancement methods. The reconstructed image quality is evaluated by the peak-signal-to-noise ratio (PSNR) and structural similarity index (SSIM) metrics by optimizing information loss in reconstructed MRI measured in mean squared error (MSE) metric.

Keywords: MRI reconstruction, Convolution Neural Network, deep learning, Rician noise, MRI denoising, image enhancement.

1. Introduction

MRI is an important and widely used non-invasive medical imaging technique in patient care which can provide high quality three-dimensional (3D) images of internal anatomical and physiological structures of human body. MRI is used in a number of clinical and research applications to provide complete structural information for disease diagnosis and prognosis of the affected human body organs. In practice, it is not always possible to acquire higher resolution high contrast MRI due to...
many real-time constraints. Sometimes a contrast medium is injected into the desired area of human body to detect and acquire high contrast images such as blood flow streaming, tumors, bone fractures and tissue cell inflammations [1]. As MRI utilizes non-ionized radiation and strong magnetic field to generate accurate image. It is safer imaging modality than other techniques such as computerized tomography and x-rays. MRI provides detailed information images of internal bone structures, soft-tissue regions and blood-flow streaming in cardiac region for diagnosis and accurate image analysis. The MRI scanners used to acquire images applying a strong magnetic field created with the scanner hardware gradient-coils generating radio waves. MRI scanners can scan the desired human body parts provide 3-D volumetric images by scanning multiple 2-D sequential slices. However, it is possible that the outer-plane resolution may lower than the in-plane direction which permits faster image acquisition and also minimize the noise levels in each 2-D slice. MR image scan takes a longer scan time to acquire high resolution and high signal-to-noise ratio (SNR). Considering patient condition and comfort, it is not preferred to scan again and again the same data. Hence, low resolution images with non-uniform intensity region gets acquired. To enhance these low resolution and non-uniform intensity region images, post-acquisition image enhancement techniques are used. This is one of the most important steps in MRI quality enhancement and reconstruction of high-resolution images. This operation aims to identify and detect the pixels groups in the image to improve their spatial contrast in the low-contrast regions. MRI signal is acquired in the image scanner in complex frequency domain by forward Fourier transform and inverse transformation. This process provides frequency spectrum of the acquired complex MR signal, called \( k \)-space which is degraded by various environmental factors such as random additive white Gaussian noise, body motion artifacts, scanner electric voltage fluctuations etc. When the spatial 2-D computer readable digital image is constructed taking squared magnitude of real-part and imaginary-part of the complex data, the resulted image is created with low intensity regions and degraded quality with noise artifacts [2]. The MR image enhancement process is the transform of image data into improved perceptual quality of the given image which enable the observers to see the image information which is not be immediately available in the original acquired image. This can happen when the dynamic intensity range in the image of that not commensurate with display device when image contains high noise contents or having low contrast. Main motive of the image enhancement technique is to provide high perceptual quality MRI with maximum structural details for visually analysis and for automated quantitative processing by computer algorithms. Rest of the paper has been organized as: Section 2 gives background of basic MRI filters used in noise removal and enhancement of spatial low intensity image regions. Section 3 elaborates the traditional frequency domain filtering techniques exploits the conditional probability of 2-D and 3-D MRI and self-similarity property of image regions. Section 4 contains proposed deep convolutional neural network filtering details for MRI. Section 5 gives experimental and analysis details and in section 6, conclusion and future scope has been given.

2. Background

Image enhancement techniques are mathematical procedures designed to manipulate and enhance the MRI contrast in the low-contrast image regions. A set of operations is performed on low contrast images to enhance its perceptual quality to make it suitable for automated clinical analysis. Conventionally, various image quality enhancement methods have been applied to low-resolution images as a post-processing operation [3]. Conventionally, histogram equalization (HE), adaptive histogram equalization (AHE) and contrast limited adaptive histogram equalization (CLAHE), image enhancement methods have been used to scale up the pixel intensities in the low-resolution MRI [4][5]. Despite the tremendous advancement in the digital signal sensing technologies, the problem of acquiring good quality MRI with high resolution is still facing many challenges. Over the last many years, medical MRI denoising reconstruction research area witnessed a commendable growth and the
process is continuously flourishing with new research in this domain. The reason behind this is that no single technique is capable in handling the issues and challenges occurred in medical MRI. A fundamental reason in MRI enhancement and reconstruction is to provide perceptually good quality images with visible fine details to the human observer for analysis. Also, MRI with high signal-to-noise ratio (SNR) found suitable for automated image analysis and for further image processing tasks like, tumor segmentation, classification, registration and edge detection. The major issues in MRI acquisition is how to balance efficiently and optimally between the MRI spatial resolution, SNR and scan-time. These three factors are highly inter-dependent. Spatial high-resolution MRI enable the observers to perceive finer image details, but it comes on the cost of low-SNR or longer scan-time. During image acquisition, it is desirable to scan image with adequate SNR level for perceptual analysis of region-of-interest in the image. Availability of MRI scanning resources is limited and costly. Long scan-times are not preferred to acquire high SNR MRI due to patient condition and comfort. Moreover, in long scan-time, patient’s body movement also induce motion artifacts in the acquired MRI. Due to these three parameters, post-acquisition image processing and enhancement techniques are preferred to reconstruct high resolution MR images [3] from the low quality as shown below:

![Figure 1](image)

**Figure 1**: Starting from left, (a) is the normal acquired image, (b) is the low intensity image and (c) is the enhanced MRI slice.

Historically, image quality enhancement started from calculating and equally distributing the intensity levels of all image pixels by histogram equalization in all pixel locations. The low contrast image regions contain pixel vectors with each intensity level count at each pixel position. The histogram $hist(i)$ for image intensity levels can be defined from total number of pixels of each intensity level as:

$$h(i) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} H(f(x,y) - i)$$  \hspace{1cm} (1)

where as operator $H$ is defined is:

$$H(i) = \begin{cases} 1, & \text{if } i = 0 \\ 0, & \text{otherwise} \end{cases}$$  \hspace{1cm} (2)
A normal image intensity distribution and redistribution after equalization of intensity histogram is shown in figure-2 below.

A local convolution filter mask or window \( W(k, l) \) of odd size dimensions is used to scale the pixel intensity by convolving the filter weights with image pixels. A kernel window can be of size \((2K + 1 \times 2L + 1)\) coefficients, where each pixel location \((x,y) = (0)\) is the center of filter window and the convolution operation can be defined as:

\[
g(x, y) = W * f(y) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} W(k,l) * f(x-k, y-l)
\]  

where \( g(x,y) \) is the output image of the convolution operation of \( W(k,l) * f(y) \) filter and input image.

Figure 2: First, left-side figure is a normal histogram and second, right-side is of intensity equalized histogram of the image shown in figure 1 above.

The convolution filter weight parameters are convolved with the input image in order to upscale the pixel intensity levels by a specific enhancement operation. This will modify the desired pixel values by amplifying their intensity while suppressing unwanted pixel intensities penalizing them with specified kernel weights. The specific values of the kernel weights can be set depending on the different types of enhancements operations. In frequency domain, the image enhancement operation is performed by Fourier transform. The Fourier transformation \( F(u,v) \) on an image \( f(x,y) \) can be performed as:

\[
F(u,v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y)e^{-2\pi ij \left( \frac{u}{M} \right) \left( \frac{v}{N} \right)}
\]  

where \( u = 0,1,2,..., M-1 \) and \( v = 0,1,2,..., N-1 \) are transformed image coefficients of the two-dimensional input image. A spatial domain 2-D image can be reconstructed from the inverse of transform domain coefficient which is given here as:
where \( f(x, y) \) is the reconstructed output image. \( x = 0, 1, 2, \ldots, M - 1 \) and \( y = 0, 1, 2, \ldots, N - 1 \) are spatial pixel locations of \( M \times N \) MRI [6].

2.1. Pixel Intensity Scaling

Image pixel intensity levels is scaled to enhance the intensity levels in the range of the desired regions of the acquired MRI which can be in a very narrow range of intensity band. The limited band intensity can be scaled focusing on specific intensity bands. For example, if two images, \( f_1(x, y) \) and \( f_2(x, y) \) are known to be defined for the pixel intensity band of interest, an intensity scaling transformation can be applied by defining the same as:

\[
\begin{align*}
\text{eim} &= \begin{cases} 
  f, & f_1 \leq f \leq f_2 \\
  0, & \text{otherwise}
\end{cases} \\
\text{and} \\
g &= \begin{cases} 
  \text{eim} - f_1, \\
  f_2 - f_1, \quad (f_{\text{max}})
\end{cases}
\end{align*}
\]

where parameter \( \text{eim} \) is an intermediate image and \( g \) is the output image with maximum pixel intensity as \( \text{asf}(\text{max}) \).

2.2. Intensity Equalization

The low-resolution image regions enhanced by stretching specific pixel intensity levels. For this purpose, the probability distribution of pixel intensities is normalized to map with a maximum flat histogram intensity levels which ranges from 0 to \( L - 1 \) gray-levels for the total pixels of \( M \times N \) size image. For equal distribution of the intensity levels over all image regions, each level of histogram represents with pixel counts of \( M \times N \) image. A very simple approach to redistribute the image pixels is the normalized cumulative histogram which is defined as:

\[
H(j) = \frac{1}{MN} \sum_{i=0}^{j} \text{hist}(i)
\]

where \( j = 0 \) to \( L - 1 \) is the maximum gray-scale intensity levels i.e. 0-255, and \( H(j) \) is the normalized cumulative histogram. Now, \( H(j) \) is used to map pixel intensity levels between the original input image and the scaled intensity gray-levels required for image quality enhancement.

The enhanced image \( \hat{f}(x, y) \) will have a maximal equal intensity distribution in all its regions, and is defined as:

\[
\hat{f}(x, y) = (L - 1) * H(f(x, y))
\]

2.3. Enhancement by noise suppression

When MRI quality is degraded with any amount of the random noise contents, traditionally local filtering method is applied to assign a new pixel value for each pixel in the image by calculating it
from the neighborhood pixels values around that pixel. These operators can be different sizes and weight values. The linear transform-based spatial domain filters are designed to convolve in a sliding-window fashion and some are non-linear. These filters are applied to update the center value under the operator by a mathematical or statistical operation such as calculating mean or median, to replace the original value of the pixel. The convolution operation is performed by convolving the filter window of \((2K+1\times2L+1)\) \((K = \text{horizontal Rows}) \times (L = \text{vertical Columns})\) size.

### 2.4. Edge Enhancement

MRI is also enhanced highlighting the edges and boundary pixels of a region within the image. Edges can be selectively identified in different orientation of image details and enhanced by a key factor on horizontal and vertical edge directions. Following edge masks can be applied for horizontal edge or line enhancement as:

\[
W_{Hz1}(k, l) = \begin{bmatrix}
1 & 1 & 1 \\
0 & 0 & 0 \\
-1 & -1 & -1 \\
\end{bmatrix} \quad \text{or} \quad W_{Hz2}(k, l) = \begin{bmatrix}
-1 & -1 & -1 \\
0 & 0 & 0 \\
1 & 1 & 1 \\
\end{bmatrix}
\]

(10)

and for vertical edge or line enhancements these filters can be defined as:

\[
W_{Vz1}(k, l) = \begin{bmatrix}
1 & 0 & -1 \\
1 & 0 & -1 \\
1 & 0 & -1 \\
\end{bmatrix} \quad \text{or} \quad W_{Vz2}(k, l) = \begin{bmatrix}
-1 & 0 & 1 \\
-1 & 0 & 1 \\
-1 & 0 & 1 \\
\end{bmatrix}
\]

(11)

A single all-directional mask, named as un-sharp mask can be applied to enhance all directional edges in an image which is defined as:

\[
W_{HP}(k, l) = \begin{bmatrix}
-1/8 & -1/8 & -1/8 \\
-1/8 & 1 & -1/8 \\
-1/8 & -1/8 & -1/8 \\
\end{bmatrix}
\]

(12)

These filters with positive filter weights produce an output image with positive and negative pixel values. A high contrast image with only positive pixel values can be reconstructed by setting an threshold criteria for selecting the desired intensity levels of image pixels.

### 2.5. Local Area Histogram Equalization

The local-region based image enhancement operation can be applied to equalize the all pixel’s intensity levels in an image by overlapping the local-region intensity levels \([7, 8]\). This nonlinear operation will significantly improve the perceptual image quality of low intensity image details. The local-area histogram of an image region intensity levels for its each low-intensity pixel location \((x, y)\) can be defined as:

\[
h_{LA}(x, y)(i) = \sum_{k=-1}^{K} \sum_{l=-1}^{L} \delta(f(x+k, y+l) - i)
\]

(13)

where \(i = 0 \text{ to } L - 1\) is the gray-scale intensity levels are representing the \(h_{LA}(x, y)(i)\)hist which is called local area cumulative histogram.
Finally, the output image \( g(x,y) \), with histogram equalized intensity level distribution is reconstructed setting filter mask of size \( K = L = 15 \) i.e. the mask window size of \( 31 \times 31 \) as:

\[
H_{LA}(x,y)(j) = \frac{1}{(2K+1) \times (2L+1)} \sum_{i=0}^{l} h_{LA}(x,y)(i)
\]  

(14)

and the equalized histogram image \( g(x,y) \) is recovered from the localized histogram as:

\[
g(x,y) = (L - 1).H_{LA}(x,y)(i)
\]  

(15)

2.6. Image enhancement and denoising by averaging

A high-quality image can be restored from the original noisy image by suppressing noise contents by averaging operation based on following three prior conditions:

(i) When relatively large number of images are available of same object.
(ii) Each image is contaminated with same nature and type of noise
(iii) Noise is independently and identically distributed with zero mean and equal variance in the image.

Based on these three priors, multiple images of the same scene or object are used to calculate average and assigned to each pixel location in the output image. This method can effectively restore and significantly enhance even severely noise corrupted images. Each \( i^{th} \) noisy input image \( f_{Noisy(i)}(x,y) \) is represented as:

\[
f_{Noisy(i)}(x,y) = f(x,y) + \eta(x,y)
\]  

(16)

Here \( f(x,y) \) is the clean image and \( \eta(x,y) \) is the random additive type noise component. If a total number of available images is \( Q \), then an averaged output image \( g(x,y) \) is computed as:

\[
g(x,y) = \frac{1}{Q} \sum_{i=1}^{Q} f_{Noisy(i)}
\]  

(17)

The expected image \( E\{g(x,y)\} \) obtained by the expectation operator \( E\{\cdot\} \) from the \( g(x,y) \) such that the expected image value \( g(x,y) \) is very close to the original input image \( f \) and standard deviation of the output image \( g(x,y) \) is taken as:

\[
\sigma_g = \frac{\sigma_d}{\sqrt{Q}}
\]  

(18)

whereas \( \sigma_d \) is the noise standard deviation in the image. When large the averaging image data-set \( Q \) will be, higher the quality of averaged noise free image.

2.7. Enhancement by Image Subtraction

MRI is also enhanced by subtraction operation between two or more images. By this method, the estimated difference equal to the noise components is subtracted from the noisy image. Images
acquired under different conditions are not required to be registered. When two images \( f_1(x, y) \) and \( f_2(x, y) \) are given, then the difference between two images \( b(x, y) = f_2(x, y) - f_2(x, y) \) is used to obtain the re-scaled output \( g(x, y) \) as:

\[
g(x, y) = f_{\text{max}} \left( \frac{b(x, y) - \min \{b(x, y)\}}{\max \{b(x, y) - \min \{b(x, y)\}\}} \right)
\]  

(19)

Whereas \( f_{\text{max}} \) represent the maximum intensity levels available in the obtained image \( g \). The \( b(x, y) \) is a un-scaled input image. The component \( \min \{b(x, y)\} \) and \( \max \{b(x, y)\} \) are the minimum and maximal intensity values respectively in the \( b(x, y) \) image.

2.8. Frequency Domain MRI Enhancement Techniques

The linear transform image enhancement methods can also be applied in frequency transform by the forward Fourier transform and inverse Fourier transform pair to the noise corrupted input image. The output image \( g(x, y) \) is obtained by convolving an input image with a sliding window intensity modifier filter mask \( W(k, l) \) which can be defined as:

\[
g(x, y) = W(k, l) * f(x, y)
\]  

(20)

The frequency transformed complex coefficients are obtained in \( G(u, v) \) from the spatial Fourier transform frequency spectrum of the input image \( g(x, y) \) as:

\[
G(u, v) = W(k, l) * F(u, v)
\]  

(21)

where \( W(k, l) \) is the filter mask and \( F(u, v) \) is the Fourier transform of basis function and is computed as defined in Eq. (4 and 5) above. A hard or soft threshold criterion is followed to modify the noise coefficients while preserving the high edge details in the image.

The threshold pixel values \( D(u, v) \) is computed from the square root of the squared sum \( \sqrt{u^2 + v^2} \) from a threshold criteria \( T \) of \( D_T \) used to determine the \( (u, v) \) pixel value. For constructing a smoothing region output image, the simplest approach is to use the ideal low-pass filter \( W_L(u, v) \) which modify the pixel value as 1 when \( D(u, v) \leq D_T \) and 0 otherwise. Similarly, ideal high-pass filter \( W_H(u, v) \) is used to update a pixel assigning 1 when \( D(u, v) \geq D_T \) and 0 otherwise. But due to inception of ringing effects in images by these filters while noise removing and smoothing, these filters are not preferred.

3. Related Work

The enhanced MRI reconstruction from the noise corrupted signal is a well-studied ill-posed research problem where no unique solution exists or many solutions exists. In computer vision and image processing tasks, various ill-posed problems exist such as image denoising, deblurring, inpainting, motion correction etc. Various MRI denoising and enhancement techniques have been presented over the years from which some of the techniques outperformed others and achieved the state-of-the-art image denoising and enhancement performance. A new wavelet-based MRI denoising and reconstruction methods has been presented in [9] by shrinking Laplacian distributed wavelet coefficients based on conditional probability of each coefficient considering it as noise or image pixels. In [10], Chang et al, proposed two-stage block-wise 3D non-local means (3DNLM) method to
restore the noisy MRI slices from neighborhood-slices and then applying multi-dimensional PCA as a post processing step to denoising MRI while preserving the important image information. In [11], the higher-order singular-value decomposition (HOSVD) algorithm presented for volumetric 3D MRI denoising which achieved comparable performance that of famous block-matching BM4D for 3D medical images. Manjon et al presented 2D nonlocal means noise filtering based new methods for Rician distributed noise filtering in [12] for 3D MRI denoising and reconstruction exploiting self-similar image regions in noise corrupted 3D MRI [13], further an extension of the self-similar patch-based sub-bank wavelet mixing for 3D MRI denoising is presented and in [14], and an optimized block-wise non-local means filter is presented for 3D MRI denoising, which outperformed other existing denoising methods at that time. Similarly, in [15] presented a two-stage noise estimation and removal method exploiting sparsity and self-similarity property of image regions for denoising MRI. First, the non-local PCA based threshold is applied to image coefficient to automatically estimate the noise contents and then rotationally invariant non-local means noise filtering applied which automatically estimate the noise in spatially varying noise levels and correct the bias effect induced by Rician probability distribution modeled image. A non-local means (NLM) along with Laplacian of Gaussian (LoG) filter has been applied in [16] on squared magnitude MRI simultaneously to filter out noise contents and compensating bias effect efficiently while preserving the structural details in the image. An artificial neural network (ANN) presented in [17] to predict the noise parameter in MRI with texture feature analysis for automatic denoising. The proposed method effectively restores the MRI and process the results very fast. A linear minimum mean square error estimation-based 3D-MRI restoration technique is presented in [18] which exploits self-similar property of MRI regions to restore the images while preserving the important structural details. The noise parameter is estimated with Bayesian mean square error to address the denoising problem.

A supervised feature learning methodology has been extended developing the unsupervised discriminative-learning models for processing large imaging databases. A multi-layer network model is used to extract the multiple features in image convolving a set of different weight filters to memorize the image details to preserve the important structural information. In [19], a feedforward denoising convolutional neural network model has been presented for natural images. This model consists many hidden layers, conforming a deep neural network architecture which can learn every important image feature from each orientation. This model exploits the random Gaussian noise affected image pixels with unknown noise levels. The difference between the input noisy image and target image is estimated by learning the difference between the input and estimated output image. The model can learn to improve from the training images and removes the noise in intermediate hidden-layers. In [19], it has been proved that deep CNN model can restore the maximum useful image information while performing denoising operations. CNN methods are more efficient to reconstruct the original image from the heavily noise corrupted images. Next in [20] a fast and flexible image denoising method has been presented which is more robust to handle the varying white Gaussian noise levels in degraded images. A robust multi-channel (MCDnCNN) deep learning CNN model has been presented in [21] with noise residual-learning for denoising Rician distributed 3-D MRI. This technique significantly denoised MRI with varying noise levels and outperformed other state-of-the-art denoising schemes. In [22] a noise residual based discriminative learning scheme is proposed with fully connected CNN model for multi-dimensional feature map extraction from noisy MRI focusing attention on data loss during each layer with varying noise levels. This scheme extracts the feature maps by normal and dilated convolution operation in parallel manner. A Bayesian shrinkage method is applied to wavelet transform image coefficients combining block-based auto-encoder network to denoise Rician distributed brain MRI in [23] which shows significant performance. Similar approach has been proposed in [24] to reconstruct a denoised MRI from its under-sampled \textit{k-space} data to speed-up the data acquisition. An un-decimated wavelet transform
coefficients are utilized as a prior to train the denoising auto-encoder network which are obtained from transformed highly redundant feature maps at multiple levels which enable robust network driven prior to learn. In [25], a deep learning based regularized scheme is presented to reduce the MRI scan-time. This method utilizes the compressive sensed acquisition of \( k \)-space MR signal from which MRI is reconstructed. A calibration-less compressed sensing regularizer is used to control the over-flow and under-flow. A stacked convolutional auto-encoder is used for noise reduction based on noise estimation map from coil-wise data sensitivity-map.

4. Methods
The deep noise-residue learning convolutional neural network (CNN) model learns the noise levels in the noisy input image blocks and subtract the residue noise from the final output image to produce a denoised and enhanced MR image. A CNN model requires a prior input image in order to learn the useful feature maps.

The primary aim of the MRI restoration is to provide a high-quality MRI with enhanced feature maps and preserving the important structural details in the clean image. Let \( g \in \mathbb{R}^{x \times y} \) be denoted as a input noisy MRI and \( \hat{f} \in \mathbb{R}^{x \times y} \)denoted as the corresponding restored clean image. The input and output correspondence between noise corrupted and denoised image is represented as:

\[
\hat{f} = \delta(g)
\]  

(22)

where \( \delta(.) \) is a noise content mapping function. In deep CNN models, the noise characteristics are independent of statistical mapping. Therefore, the denoising optimally approximated by the inverse of \( \delta^{-1} \) and noise is measured from the residue difference of input noise and output denoised clean image which can be defined as:

\[
\text{arg min}_\hat{f} \| g - \hat{f} \|_2^2
\]  

(23)

and where \( \hat{f} = \delta(g) \)is used to estimate the \( g \) and \( f \) is the optimal approximation of \( \delta^{-1} \)[26].

4.1. Network Architecture
As shown in the next Figure, the first layer of the denoising network takes input image which convoluted with the randomly initialized filter weights adding bias term followed by a rectification function ReLU is applied. The network’s output layer generates same sized denoised image as is the input image. A padding scheme is used with processing patch to have a same output size. No max-pooling or min-pooling layer is applied because it eliminates the useful features. The convolution and deconvolution layers work in a symmetric way to predict the pixel-wise denoising and the input size can be arbitrary. The first convolution layer includes 128 of \( 3 \times 3 \) size weighted filter kernels to extract and learn the feature map from single channel gray-scale images. Patch-similarity measure is performed considering the neighboring slice patches of size \( 32 \times 32 \). Further, the residual learning and batch normalization is used to generalize intermediate data and speed-up the network training process and for improving the denoising performance. At the last outer layer, a deconvolution is performed to reconstruct the noise free image. Following figure shows the proposed residue learning denoising convolution neural network architecture. This is a fully convolution-deconvolution network model without pooling layer to learn the all feature details in the image. The network has been trained with Adam optimizer to control and avoid the over-fitting and under-fitting problem. The learning rate is fixed at \( 10^{-6} \) and standard deviation of the noise is used maximum up to 70 percentage.
4.2. MRI Dataset
A 3D simulated BrainWeb [27] MRI phantom database has been used to evaluate the denoising performance of the proposed model. This is a freely available MRI phantom database called BrainWeb [27], which comprises of (a) T1-W, (b) T2-W and (c) Proton Density (PD) each having resolution of 181 × 217 × 181 pixels.

![Figure 3: CNN based residual MRI denoising.](image)

4.3. Network Training
A major constraint in deep learning CNN models is the non-availability of huge amount medical MRI data for training purposes. A single subject huge amount of MRI data is normally not available for training of medical MRI denoising tasks. The proposed model trained on 100 MRI slices each from T1-W, T2-W and PD volumes from BrainWeb database. The MRI database partitioned into three subsets i.e. training set, testing and validation sets. 70 slices used for training purpose and 15 from each modality. In these three types of 70 slices, synthetic noise is added in the range of 1 to 25. Further these noise corrupted data are fed to the proposed network model to perform denoising task. Now suppose $g$ is the input image, the convolution and deconvolution of the input image can be represented as:

$$
\text{Conv} \times \text{ReLU} \times \text{Deconv}$$
\[ R(l) = \max (0, W_i \ast l + B_i) \]  

(24)

where \( W_i \) and \( B_i \) are denoising filter weights and bias terms respectively. Operator \( \ast \) indicating the convolution and deconvolution operations. The pixel-wise convolutional sum of the two input images after applying the ReLU is represented as:

\[ R(g_1, g_2) = \max (0; g_1 + g_2) \]  

(25)

For end-to-end feature mapping between input noisy and output filtered image, convolutional and deconvolution network filter weights are continuously updated for each convolution and deconvolution pair \( \{g^i, f^i\} \) which are noise corrupted \( g^i(x, y) \) and \( f(x, y) \) is the original ground-truth noise free images respectively.

### 4.4. Loss Function

The information loss in image restoration and reconstruction process usually calculated by aggregating the squared difference of mean values between the noisy image and corresponding reconstructed denoised image [19, 28]. The mean square error (MSE) metric is used to compute information loss \( L \) for each image patch-pair to tune the network weights computing the loss after each convolution-deconvolution process as:

\[ L = \frac{1}{MN} \sum_{i,j=0}^{M-1,N-1} \| R(g^i, \theta) - f^i \|_F^2 \]  

(26)

where \( R \) is the noise residue difference between the denoised recovered noise free image \( f^i \) and ground-truth real noise free image \( f^i \) of \( M \times N \) pixel square size. The network learns feature maps from the input noisy image \( g(x, y) \). It has been observed that optimization for the noisy image better converges than denoised image. Empirically, optimizer Adam [29] with base learning rate \( 10^{-4} \) has been found fast converging than SGD [30]. The learning rate is set same for all layers. The gradients \( \nabla \) with respect to network filter weights of \( l^{th} \) layer is computed as:

\[ G = \nabla \theta_i L(\theta_i) \]  

(27)

### 4.5. Back-propagation to gradients update

Back propagation is used to feed backward the gradients information from the connected layers. After two convolution \( C_1 \) and \( C_2 \) of first layer input \( f_1 \), the output is \( g_1 \). The filter weight parameters \( \theta_2 \) is update by deriving the information loss from these convolution and deconvolution as:

\[ \nabla \theta_2 L(\theta_2) = \frac{\partial L}{\partial f_1} \frac{\partial f_1}{\partial \theta_2} + \frac{\partial L}{\partial f_2} \frac{\partial f_2}{\partial \theta_2} \]  

(28)

where \( f_1 \) and \( f_2 \) are input image blocks and these are similar for further layered inputs.

After obtaining noise residue for every MRI slice, the entire noise free MRI volume can be restored. Conventionally, first the pixelwise noise variance is estimated statistically from the image noise characteristics. This approach affects the denoising performance according to estimation.
accuracy. The deep learning-based CNN models demonstrate the robustness in general image denoising. This CNN model is first trained with a specific noise value to evaluate its performance with other existing denoising algorithms. Further, this model extended to single channel gray-scale Rician probability distributed MRI with unknown noise variance.

5. Experimental Evaluation and Analysis

The proposed method has been tested using freely available MRI phantom database BrainWeb [27] comprising T1-W (weighted), T2-W and Proton Density (PD) 3D MRI volumes, each having resolution of 181×217×181. The performance and results are analyzed comparing with other existing methods such classical non-local means MRI denoising [12], [14], [13] by adding new features such as fully automated smoothing parameter tuning, selection of relevant blockwise image voxels and parallel processing into denoising filters. This method shows improved performance than non-local means. Similarly, 3D wavelet sub-band mixing noise removal used while keeping the computation time low. 3D oracle-based DCT (ODCT3D) which is rotationally invariant denoising filter and an extension of non-local means for 3D MRI. A pre-filtered MRI with DCT based on threshold noise suppression is filtered by this method. One another denoising filter which is call NLMPCA [15]. The model is evaluated with known and unknown Rician probability distribution noise in MRI with varying level from 3% to 15% with patch-size 32×32 using sliding window scheme to extract the MRI patches to train the corresponding noise filter kernels.

The denoising performance of the CNN model is evaluated measuring its quantitative and qualitative scores based on two well-known performance indicators metrics, peak-signal-to-noise ratio (PSNR) and structural similarity index (SSIM) [31-35]. These metrics are defined as:

$$PSNR = 10 \log_{10} \left( \frac{255^2}{MSE} \right)$$  \hspace{1cm} (29)

where $MSE$ is computed as information loss as sum of the mean squared error between ground-truth noise free image and restored denoised image as given in Eq. (26) above. The following Table-1 shows the practical simulation results of the proposed MRI denoising method and comparison with other existing state-of-the-art denoising methods.

| Noise Levels | 3% | 5% | 7% | 9% | 11% | 13% | 15% | 17% | 19% | 21% |
|--------------|----|----|----|----|-----|-----|-----|-----|-----|-----|
| Methods      |     |    |    |    |     |     |     |     |     |     |
| PRINLM       | 37.8 | 34.8 | 32.8 | 31.3 | 30.1 | 29.0 | 28.0 | 27.1 | 26.3 | 25.5 |
| B3DNLM       | 38.0 | 35.1 | 33.1 | 32.5 | 30.1 | 29.0 | 27.9 | 27.0 | 26.2 | 25.3 |
| FFDNet       | 37.6 | 34.8 | 32.9 | 31.6 | 30.5 | 29.6 | 28.8 | 28.0 | 27.4 | 26.7 |
| MCDnCN       | 38.6 | 35.5 | 33.5 | 32.1 | 30.9 | 30.0 | 29.0 | 28.2 | 27.5 | 26.8 |
| CNNDMR       | 39.6 | 36.5 | 34.3 | 32.7 | 31.5 | 30.5 | 29.6 | 28.9 | 28.2 | 27.6 |

Table 1: Comparison of the PSNR (dB) values of the existing vs proposed MRI denoising method.
Figure 5: Starting from left-top of first row (a) is the reference noise T1-W, (b) is noisy with 9% Rician noise and (c) is denoising recovered MRI. Similarly, in second and third rows (d) to (e) are reference, 9% noisy and denoised T2-W and (g) to (i) are reference, 9% noisy and denoised PD MRI slices.
The structural information restoration in the recovered image is measured by structural-similarity score between the real noise-free image and denoised recovered image which is defined as:

$$SSIM(f, f') = \frac{(2\mu_f \mu'_f + C_1) \times (\sigma_f \sigma'_f + C_2)}{\mu^2_f + \mu^2'_f + C_1} \times \frac{(\sigma_f \sigma'_f + C_2)}{\sigma^2_f + \sigma^2'_f + C_2}$$  \hspace{1cm} (30)$$

where $\mu_f$ and $\mu'_f$ are means of reference noise-free image and recovered denoised images; $C_1$ and $C_2$ are constant variables; $\sigma_f$ and $\sigma'_f$ are variances and $\sigma_f$ and $\sigma'_f$ are the covariance of both the images.

![Noise Levels vs PSNR (db) Score in 3D MRI Denoising](image)

**Figure 6**: Noise levels present in the noisy MRI vs PSNR (db) values. Comparison of the existing and proposed MRI denoising methods.
Figure 7: Structural Similarity Index score calculated with existing MRI denoising and proposed methods.

Table-2 next shows the image reconstruction Structural Similarity Index (SSIM) performance results between the reference image and denoised images and their comparison with other existing MRI denoising enhancement methods.

| Noise Levels | 3%  | 5%  | 7%  | 9%  | 11% | 13% | 15% | 17% | 19% | 21% |
|--------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Methods      |     |     |     |     |     |     |     |     |     |     |
| PRINLM       | 0.97 | 0.95 | 0.92 | 0.89 | 0.86 | 0.83 | 0.79 | 0.76 | 0.72 | 0.69 |
| B3DNLM       | 0.97 | 0.95 | 0.93 | 0.90 | 0.87 | 0.83 | 0.79 | 0.76 | 0.72 | 0.68 |
| FFDNET       | 0.97 | 0.95 | 0.33 | 0.91 | 0.89 | 0.88 | 0.86 | 0.84 | 0.83 | 0.81 |
| MCDnCNN      | .098 | 0.96 | 0.94 | 0.92 | 0.90 | 0.89 | 0.87 | 0.85 | 0.84 | 0.82 |
| RLCNNMR I    | **0.987** | **0.974** | **0.959** | **0.941** | **0.923** | **0.905** | **0.890** | **0.873** | **0.863** | **0.846** |

Table 2: Structural Similarity (SSIM) index of the existing and proposed MRI denoising method

6. Conclusion
In this paper, a deep learning data-dependent medical image denoising technique is presented to denoise MRI using convolutional neural network. This model learns the image features from the residual noise maps extracted by subtracted each residue patch from the noisy image patch. The
The proposed method is based on specific noise feed-forward model which outperform some existing denoising methods to provide a high-quality MRI with respect to PSNR and SSIM quality metrics. We have used only single BrainWeb MRI database and further required to evaluate with other real clinical and other medical imaging data sets. An MRI denoising and enhancement method has been presented based on convolution, rectification, batch-normalization and deconvolution processing to generalize the weighted filtered images. The CNN-based image processing and enhancement methods shown significant performance over the traditional techniques. With the advancement in powerful computing resources and availability of large volume cheaper storage device, it becomes possible to mine the voluminous imaging databases. Therefore, the emerging deep learning-based CNN model learn the important feature details from the large imaging data sets and trained to enhance the image quality by down-sampling and up-sampling through the encoding-decoding process. The network trained on the BrainWeb MRI phantom to memorize the detailed feature maps from varying noise corrupted volumes. Testing and validation is performed on the noise degraded and noise free MRI and demonstrated promising quantitative and perceptual improvement in the results. In the conclusion, the proposed deep convolution network-based MRI denoising method has provided better denoising results in 2D and 3D MRI corrupted with Rician distributed noise.

References

[1] Kaur H and Rani J 2016 Mri brain image enhancement using histogram equalization techniques 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET) (IEEE) pp 770–773
[2] Bankman I 2008 Handbook of medical image processing and analysis (Elsevier)
[3] Plenge E, Poot D H, Bernsen M, Kotek G, Houston G, Wielopolski P, van der Weerd L, Niessen W J andMeijering E 2012 Magnetic resonance in medicine 68 1983-1993.
[4] Toenies K D 2017 Guide to medical image analysis (Springer)
[5] Kalyani J and Chakraborty M 2020 Contrast enhancement of mri images using histogram equalization techniques 2020 International Conference on Computer, Electrical Communication Engineering (ICCECE) pp 1–5
[6] Gonzalez R C, Woods R E et al. 2002 Digital image processing
[7] Ketcham D J, Lowe R and Weber J 1976 Real-time image enhancement techniques Seminar on Image Processing vol 74 (Pacific Grove California, Hughes Aircraft Company) pp 1–6
[8] Pizer S M, Amburn E P, Austin J D, Cromartie R, Geselowitz A, Greer T, terHaarRomeny B, Zimmerman J B andZuiderveld K 1987 Computer vision, graphics, and image processing 39 355–368
[9] Martin-Fernandez M and Villullas S 2015 Computational and mathematical methods in medicine 2015 [10] Chang L, ChaoBang G and Xi Y 2015 Computational and mathematical methods in medicine 2015 [11] Zhang X, Xu Z, Jia N, Yang W, Feng Q, Chen W and Feng Y 2015 Medical image analysis 19 75–86
[10] Kishor Kumar Reddy C, Anisha P R, Apoorva K, “Early Prediction of Pneumonia using Convolutional Neural Network and X-Ray Images”, Smart Innovation, Systems and Technologies, 2021
[11] Kishor Kumar Reddy C, Anisha P R, Shastry R, Ramana Murthy B V, “Comparative Study on Internet of Things: Enablers and Constraints”, Advances in Intelligent Systems and Computing, 2021
[12] Coupé P, Yger P, Prima S, Hellier P, Kervrann C and Barillot C 2008 IEEE transactions on medical imaging 27 425–441
[13] Manj´on J V, Coup´e P, Buades A, Collins D L and Robles M 2012 Medical image analysis 16 18–27
[14] Coup´e P, Hellier P, Prima S, Kervrann C and Barillot C 2008 International Journal of Biomedical Imaging 2008
[15] Manj´on J V, Coup´e P and Buades A 2015 Medical image analysis 22 35–47
[16] Bhujle H V and Chaudhuri S 2013 Magnetic resonance imaging 31 1599–1610
[17] Chang Y N and Chang H H 2011 Bio-medical materials and engineering 26 S1275–S1282
[18] Golshan H M and Hasanzadeh R P 2014 IEEE/ACM transactions on computational biology and bioinformatics 12 861–870
[19] Zhang K, Zuo W, Chen Y, Meng D and Zhang L 2017 IEEE transactions on image processing 26 3142–3155
[20] Zhang K, Zuo W and Zhang L 2018 IEEE Transactions on Image Processing
[21] Jiang D, Dou W, Vosters L, Xu X, Sun Y and Tan T 2018 Japanese journal of radiology 36 566–574
[22] Aetesam H and Maji S K 2021 Biomedical Signal Processing and Control 66 102405
[23] Juneja M, Saini S K, Kaul S, Acharjee R, Thakur N and Jindal P 2021 Biomedical Signal Processing and Control 69 102844
[24] Wang S, Lv J, He Z, Liang D, Chen Y, Zhang M and Liu Q 2021 Neurocomputing 437 325–338
[25] Islam S R, Maity S P and Ray A K 2021 Biomedical Signal Processing and Control 66 102399
[26] Ran M, Hu J, Chen Y, Chen H, Sun H, Zhou J and Zhang Y 2019 Medical image analysis 55 165–180
[27] Cocosco, Kollokian, Kwan, Evans accessed June 10, 2021 URL https://brainweb.bic.mni.mcgill.ca/brainweb/
[28] Lin G, Wu Q, Qiu L and Huang X 2018 Neurocomputing 275 1219–1230
[29] Kingma D P and Ba J 2014 arXiv preprint arXiv:1412.6980
[30] Dong C, Loy C C, He K and Tang X 2015 IEEE transactions on pattern analysis and machine intelligence 38 295–307
[31] Eun D i, Jang R, Ha W S, Lee H, Jung S C and Kim N 2020 Scientific Reports 10 1-17
[32] Manimala M, Naidu C D and Prasad M G 2021 Wireless personal communications 116 491-511
[33] Koonjoo N, Zhu B, Bagnall G C, Bhutto D and Rosen M 2021 Scientific reports 11 1-16
[34] Freedman J N, Gurney-Champion O J, Nill S, Shiarli A M, Bainbridge H E, Mandeville H C, Koh D M, McDonald F, Kachelrie M, Oelfke U et al. 2021 Radiotherapy and Oncology 159 209-217
[35] Islam R and Islam M S 2021 Convolutional neural network-based enhancement method of 3d mri 2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST) pp 120-124